.

Apontamentos para as Aulas Teóricas de Álgebra Linear

para

LEIC-A

Nuno Martins

Departamento de Matemática

Instituto Superior Técnico

Dezembro de 2017

Índice

Matrizes: operações e suas propriedades

Definição 1. (i) Sejam $m, n \in \mathbb{N}$. Uma **matriz** A, do tipo $m \times n$ (lê-se m por n), é uma tabela de $m \times n$ números dispostos em m linhas e n colunas:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Usa-se também a notação $A = (a_{ij})_{m\times n}$ ou simplesmente $A = (a_{ij})$, na qual a_{ij} é a **entrada** (i, j) da matriz A. Se $m = n$, diz-se que A é uma **matriz quadrada** do tipo $n \times n$ (ou de ordem n) e as entradas a_{11}, a_{22}, ..., a_{nn} formam a chamada **diagonal principal** de A. Se $m \neq n$, diz-se que A é uma **matriz rectangular**.

(ii) A **matriz linha** i de A é: $\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$, para $i = 1, ..., m$. A **matriz coluna** j de A é:

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

para $j = 1, ..., n$.

(iii) À matriz do tipo $m \times n$ cujas entradas são todas iguais a zero, chama-se **matriz nula** e representa-se por $\mathbf{0}_{m\times n}$ ou simplesmente por $\mathbf{0}$. Por exemplo

$$\mathbf{0}_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{e} \quad \mathbf{0}_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

(iv) À matriz do tipo $n \times n$

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

tal que $a_{ij} = 0$ se $i \neq j$ para todos os i, j, isto é, à matriz cujas entradas fora da diagonal principal são todas nulas, chama-se **matriz diagonal**.

(v) À matriz (a_{ij}) do tipo $n \times n$ tal que $a_{ii} = 1$ para todo o $i = 1, ..., n$, e $a_{ij} = 0$ se $i \neq j$:

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

chama-se **matriz identidade** e representa-se por $I_{n\times n}$ ou simplesmente por I.

(vi) À matriz do tipo $n \times n$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

cujas entradas por baixo da diagonal principal são todas nulas, isto é, tais que $a_{ij} = 0$ se $i > j$, chama-se **matriz triangular superior**. À matriz do tipo $n \times n$

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

cujas entradas por cima da diagonal principal são todas nulas, isto é, tais que $a_{ij} = 0$ se $i < j$, chama-se **matriz triangular inferior**.

Uma matriz diz-se **triangular** se fôr triangular superior ou triangular inferior.

Exemplo 1. As matrizes

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & -2 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 7 \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$$

são dos seguintes tipos: A é 2×2, B é 2×4, C é 1×3, D é 4×1. Tem-se, por exemplo, $a_{21} = -2$, $b_{13} = 3$, $c_{12} = 0$ e $d_{41} = 1$.

Observação 1. Uma matriz (real) A do tipo $m \times n$ é uma aplicação:

$$\begin{array}{rcl} A : \{1, ..., m\} \times \{1, ..., n\} & \longrightarrow & \mathbb{R} \\ (i, j) & \longrightarrow & a_{ij} \end{array}$$

Notação 1. O conjunto de todas as matrizes reais (complexas) do tipo $m \times n$ é denotado por $\mathcal{M}_{m\times n}(\mathbb{R})$ ($\mathcal{M}_{m\times n}(\mathbb{C})$). Tem-se $\mathcal{M}_{m\times n}(\mathbb{R}) \subset \mathcal{M}_{m\times n}(\mathbb{C})$.

Definição 2. Duas matrizes são iguais se forem do mesmo tipo e se as entradas correspondentes forem iguais, isto é, $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{p\times q}$ são **iguais** se $m = p$, $n = q$ e $a_{ij} = b_{ij}$, para $i = 1, ..., m$ e $j = 1, ..., n$.

Definição 3. A **soma** de duas matrizes do mesmo tipo

$$A = (a_{ij})_{m\times n} \quad \text{e} \quad B = (b_{ij})_{m\times n}$$

é a matriz

$$A + B = (a_{ij} + b_{ij})_{m\times n}.$$

Exemplo 2. Sejam

$$A = \begin{bmatrix} 1 & 4 & -1 \\ -3 & 2 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -2 & 4 \\ 7 & 3 & 9 \end{bmatrix}, \quad C = \begin{bmatrix} -1 \\ 1/2 \\ -\sqrt{2} \end{bmatrix} \quad \text{e} \quad D = \begin{bmatrix} 1 \\ -1/2 \\ \sqrt{2} \end{bmatrix}.$$

$A + B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, $C + D = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ e não é possível, por exemplo, somar B com C.

Definição 4. O **produto de um escalar** (número real ou complexo) α **por uma matriz** $A = (a_{ij})_{m \times n}$ é a matriz:

$$\alpha A = (\alpha a_{ij})_{m \times n}.$$

Notação 2. A matriz $(-1)A$ será denotada por $-A$.

Exemplo 3. Seja $A = \begin{bmatrix} 1 & 4 & -1 \\ -3 & 2 & 6 \end{bmatrix}$. Tem-se, por exemplo, $-2A = \begin{bmatrix} -2 & -8 & 2 \\ 6 & -4 & -12 \end{bmatrix}$.

Observação 2. $1A = A, \qquad 0A = \mathbf{0}$ (**matriz nula**).

Definição 5. A **diferença** entre duas matrizes A e B do mesmo tipo é definida por

$$A - B = A + (-B),$$

ou seja, é a soma de A com o simétrico de B.

Definição 6. **(i)** O **produto** AB de duas matrizes A e B só pode ser efectuado se o número de colunas da 1^a matriz, A, fôr igual ao número de linhas da 2^a matriz, B. Nesse caso, o produto AB de $A = (a_{ij})_{m \times p}$ por $B = (b_{ij})_{p \times n}$ é definido por:

$$AB = (a_{i1}b_{1j} + ... + a_{ip}b_{pj})_{m \times n} = \left(\sum_{k=1}^{p} a_{ik}b_{kj} \right)_{m \times n},$$

isto é,

$$\begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \cdots & \vdots \\ a_{i1} & \cdots & a_{ip} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ b_{p1} & \cdots & b_{pj} & \cdots & b_{pn} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{p} a_{1k}b_{k1} & \cdots & \sum_{k=1}^{p} a_{1k}b_{kn} \\ \cdots & \sum_{k=1}^{p} a_{ik}b_{kj} & \cdots \\ \sum_{k=1}^{p} a_{mk}b_{k1} & \cdots & \sum_{k=1}^{p} a_{mk}b_{kn} \end{bmatrix}$$

Note que sendo $\mathbf{b}_1, ..., \mathbf{b}_n$ as colunas da matriz B, então

$$AB = A \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_n \end{bmatrix}$$

e sendo $\mathbf{a}_1, ..., \mathbf{a}_m$ as linhas da matriz A, então

$$AB = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} B = \begin{bmatrix} \mathbf{a}_1 B \\ \vdots \\ \mathbf{a}_m B \end{bmatrix}$$

(ii) Sejam A uma matriz do tipo $n \times n$ e $p \in \mathbb{N}$. A **potência** p de A é definida por

$$A^p = \underbrace{A...A}_{p \text{ vezes}} \quad \text{e} \quad \text{para } p = 0 \quad \text{define-se (se } A \text{ fôr não nula) } A^0 = I.$$

(iii) Diz-se que duas matrizes A e B **comutam** se $AB = BA$.

Exemplo 4. (i)

$$\begin{bmatrix} 0 & -2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ -3 & 2 & -2 \end{bmatrix} =$$

$$= \begin{bmatrix} 0 \times 1 + (-2) \times (-3) & 0 \times 1 + (-2) \times 2 & 0 \times (-1) + (-2) \times (-2) \\ 2 \times 1 + 3 \times (-3) & 2 \times 1 + 3 \times 2 & 2 \times (-1) + 3 \times (-2) \end{bmatrix} =$$

$$= \begin{bmatrix} 6 & -4 & 4 \\ -7 & 8 & -8 \end{bmatrix}$$

(ii) $\begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 1/2 \\ -\sqrt{2} \end{bmatrix} = \left[1 \times (-1) + 1 \times \frac{1}{2} + (-1) \times \left(-\sqrt{2}\right)\right] = \left[\sqrt{2} - \frac{1}{2}\right]$

(iii)

$$\begin{bmatrix} -1 \\ 1/2 \\ -\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} =$$

$$= \begin{bmatrix} (-1) \times 1 & (-1) \times 1 & (-1) \times (-1) \\ \frac{1}{2} \times 1 & \frac{1}{2} \times 1 & \frac{1}{2} \times (-1) \\ \left(-\sqrt{2}\right) \times 1 & \left(-\sqrt{2}\right) \times 1 & \left(-\sqrt{2}\right) \times (-1) \end{bmatrix} = \begin{bmatrix} -1 & -1 & 1 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\sqrt{2} & -\sqrt{2} & \sqrt{2} \end{bmatrix}$$

(iv) $p \in \mathbb{N}$, $\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}^p = \begin{bmatrix} (a_{11})^p & 0 & \cdots & 0 \\ 0 & (a_{22})^p & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & (a_{nn})^p \end{bmatrix}.$

Observação 3. **(i)** O produto de matrizes não é comutativo. Por exemplo, para

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ e } B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \text{ tem-se } AB = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ e } BA = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Logo $AB \neq BA$.

(ii) $CD = \mathbf{0} \not\Rightarrow (C = \mathbf{0}$ ou $D = \mathbf{0})$, pois, por exemplo, para

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad \text{e} \qquad D = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad CD = \mathbf{0}.$$

(iii) Se A (B) tem uma linha (coluna) nula então AB tem uma linha (coluna) nula.

(iv) MUITO IMPORTANTE: Sendo

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \ X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

então:

$$AX = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + ... + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n.$$

Definição 7. A **transposta** de uma matriz $A = (a_{ij})_{m\times n}$ é a matriz $A^T = (a_{ji})_{n\times m}$, isto é

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \cdots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.$$

Definição 8. Sendo $A = (a_{ij})_{n\times n}$ uma matriz quadrada, chama-se **traço** de A ao número real (ou complexo)

$$\operatorname{tr}(A) = a_{11} + ... + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Exemplo 5. **(i)** $\begin{bmatrix} 1 & -3 \\ 4 & 2 \\ -1 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 & -1 \\ -3 & 2 & 6 \end{bmatrix}$. **(ii)** $\operatorname{tr}\left(\begin{bmatrix} 1 & 4 \\ -3 & -2 \end{bmatrix}\right) = -1$.

Teorema 1. Sejam A, B, C e D matrizes de tipos apropriados, α e β escalares. São válidas as seguintes propriedades para as operações matriciais.

(a) (Comutatividade da soma) $A + B = B + A$.

(b) (Associatividade da soma) $A + (B + C) = (A + B) + C$. Note que esta propriedade permite generalizar a definição de **soma** de 2 matrizes à **soma** de um n^o finito de matrizes, desde que as matrizes intervenientes sejam de tipos apropriados.

(c) (Elemento neutro da soma) Existe uma única matriz $\mathbf{0}$ do tipo $m \times n$ tal que $A + \mathbf{0} = \mathbf{0} + A = A$, para toda a matriz A do tipo $m \times n$.

(d) (Simétrico) Para cada matriz A existe uma única matriz B tal que $A + B = B + A = \mathbf{0}$. Esta matriz B denota-se por $-A$.

(e) (Associatividade do produto por escalares) $\alpha\,(\beta A) = (\alpha\beta)\,A$.

(f) (Distributividade) $(\alpha + \beta)\,A = \alpha A + \beta A$.

(g) (Distributividade) $\alpha\,(A + B) = \alpha A + \alpha B$.

(h) (Associatividade do produto de matrizes) $A\,(BC) = (AB)\,C$. Note que esta propriedade permite generalizar a definição de **produto** de 2 matrizes ao **produto** de um n^o finito de matrizes, desde que as matrizes intervenientes sejam de tipos apropriados.

(i) (Distributividade) $A\,(B + C) = AB + AC \quad$ e $\quad (B + C)\,D = BD + CD$.

(j) $\alpha\,(AB) = (\alpha A)\,B = A\,(\alpha B)$. $\qquad \underbrace{A + ... + A}_{p \text{ vezes}} = pA$. $\qquad (A^p)^q = A^{pq}$.

(k) $AI = A$ e $IB = B$, para todas as matrizes $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{n\times m}$, onde I é a matriz identidade do tipo $n \times n$.

(l) $A\mathbf{0} = \mathbf{0}$ e $\mathbf{0}B = \mathbf{0}$, para todas as matrizes $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{n\times m}$, onde $\mathbf{0}$ é a matriz nula do tipo $n \times n$.

(m) $\left(A^T\right)^T = A$.

(n) $(A + B)^T = A^T + B^T$.

(o) $(A_1 + A_2 + ... + A_n)^T = A_1^T + A_2^T + ... + A_n^T$, com A_1, A_2, ..., A_n matrizes de tipos apropriados.

(p) $(\alpha A)^T = \alpha A^T$.

(q) $(AB)^T = B^T A^T$.

(r) $(A_1 A_2 ... A_n)^T = A_n^T ... A_2^T A_1^T$, com A_1, A_2, ..., A_n matrizes de tipos apropriados.

(s) Sendo $A = (a_{ij})_{n\times n}$ e $B = (b_{ij})_{n\times n}$ duas matrizes quadradas e α um escalar, tem-se

$$\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B), \qquad \mathrm{tr}(\alpha A) = \alpha\,\mathrm{tr}(A), \qquad \mathrm{tr}(A^T) = \mathrm{tr}(A) \qquad \text{e} \qquad \mathrm{tr}(AB) = \mathrm{tr}(BA).$$

Resolução de sistemas de equações lineares

Definição 9. Uma **equação linear** com n incógnitas $x_1, x_2, ..., x_n$ é uma equação da forma

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b,$$

em que $a_1, a_2, ..., a_n$ e b são constantes (reais ou complexas). A b chama-se **termo independente**.

Definição 10. Um **sistema de m equações lineares** com n incógnitas $x_1, x_2, ..., x_n$ é um conjunto de equações da forma

$$(*) \left\{ \begin{array}{c} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{array} \right.$$

em que a_{ij} e b_k são constantes (reais ou complexas), para $i, k = 1, ..., m$ e $j = 1, ..., n$.

Definição 11. Uma **solução** (caso exista) de um sistema de m equações lineares com n incógnitas reais, é o elemento

$$(s_1, s_2, ..., s_n) \in \mathbb{R}^n := \{(a_1, a_2, ..., a_n) : a_1, a_2, ..., a_n \in \mathbb{R}\}$$

que satisfaz as equações desse sistema quando substituímos

$$x_1 = s_1, \quad x_2 = s_2, \quad ..., \quad x_n = s_n.$$

(No caso das variáveis serem complexas ter-se-ia soluções em $\mathbb{C}^n$.)

Usando o produto de matrizes, isso equivale a dizer que

$$S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

satisfaz a equação matricial

$$AX = B,$$

em que

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

isto é, fazendo $X = S$ tem-se a condição verdadeira $AS = B$. Ao conjunto de todas as soluções do sistema chama-se **conjunto solução** ou **solução geral** do sistema.

Definição 12. A matriz A é a **matriz dos coeficientes do sistema** $AX = B$, X é a matriz coluna das incógnitas e B é a matriz coluna dos termos independentes. A matriz

$$[A \mid B] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & | & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & | & b_m \end{bmatrix}$$

associada ao sistema $(*)$ chama-se **matriz aumentada do sistema**.

Exemplo 6. O sistema linear de duas equações e duas incógnitas

$$\begin{cases} x + 2y = 1 \\ 2x + y = 0 \end{cases}$$

pode ser escrito do seguinte modo:

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

A solução geral do sistema acima é dada por

$$\{(x, y) : x + 2y = 1 \text{ e } 2x + y = 0\} = \{(-1/3, 2/3)\},$$

isto é, $X = \begin{bmatrix} -1/3 \\ 2/3 \end{bmatrix}$ é a única matriz que satisfaz $AX = B$, com $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ e $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Definição 13. A um sistema de equações lineares da forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = 0 \\ ... \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = 0 \end{cases}$$

chama-se **sistema linear homogéneo**. Este sistema pode ser escrito na forma $AX = \mathbf{0}$.

Observação 4. **(i)** Todo o sistema linear homogéneo $AX = \mathbf{0}$ admite pelo menos a **solução trivial**:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Assim, todo o sistema linear homogéneo tem solução. Além disso, como iremos ver, ou tem apenas a solução trivial ou tem um número infinito de soluções.

(ii) Num próximo capítulo, à solução geral do sistema linear homogéneo $AX = \mathbf{0}$ dar-se-á o nome de **núcleo** de A e escrever-se-á $\mathcal{N}(A)$.

Definição 14. Às seguintes operações que se podem aplicar às equações de um sistema de equações lineares, chamam-se **operações elementares.**

(a) Trocar a posição de duas equações do sistema;

(b) Multiplicar uma equação por um escalar diferente de zero;

(c) Substituição de uma equação pela sua soma com um múltiplo escalar de outra equação.

Definição 15. Dois sistemas de equações lineares que se obtêm um do outro através de um número finito de operações elementares, dizem-se **equivalentes**, tendo assim o mesmo conjunto solução.

Observação 5. Quando aplicamos operações elementares às equações de um sistema de equações lineares, só os coeficientes e os termos independentes do sistema são alterados. Logo, aplicar as operações elementares anteriores às equações de um sistema linear $(*)$ equivale a aplicar às linhas da matriz aumentada

$$[A \mid B] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & | & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & | & b_m \end{bmatrix}$$

as seguintes operações.

Definição 16. As **operações elementares** que podem ser aplicadas às linhas (i e j) de uma matriz são:

(i) Trocar a posição de duas linhas (i e j) da matriz: $\quad L_i \leftrightarrow L_j$

(ii) Multiplicar uma linha (i) da matriz por um escalar (α) diferente de zero: $\quad \alpha L_i \to L_i$

(iii) Substituição de uma linha (j) pela sua soma com um múltiplo escalar (α) de outra linha (i): $\quad \alpha L_i + L_j \to L_j$

Teorema 2. Se dois sistemas lineares $AX = B$ e $CX = D$ são tais que a matriz aumentada $[C \mid D]$ é obtida de $[A \mid B]$ através de uma ou mais operações elementares, então os dois sistemas são equivalentes.

Definição 17. Uma **matriz** $A = (a_{ij})_{m \times n}$ diz-se **em escada de linhas** se:

(i) Todas as linhas nulas (formadas inteiramente por zeros) estão por baixo das linhas não nulas;

(ii) Por baixo (e na mesma coluna) do primeiro elemento não nulo de cada linha e por baixo dos elementos nulos anteriores da mesma linha, todas as entradas são nulas. Esse primeiro elemento não nulo de cada linha tem o nome de **pivot**.

Exemplo 7. As seguintes matrizes estão em escada de linhas:

$$A_1 = \begin{bmatrix} 4 & -1 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & -5 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & -1 & 1/2 & 0 & 0 \\ 0 & 0 & -3 & 0 & \sqrt{2} \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Definição 18. O método de resolver sistemas de equações lineares que consiste em aplicar operações elementares às linhas da matriz aumentada do respectivo sistema de modo a que essa matriz fique em escada de linhas, chama-se **método de eliminação de Gauss**.

Definição 19. Um sistema de equações lineares diz-se:
(i) impossível se não tiver soluções;
(ii) possível e indeterminado se tiver um número infinito de soluções;
(iii) possível e determinado se tiver uma única solução.

Definição 20. (i) O número de **incógnitas livres** (podem tomar valores arbitrários) de um sistema, é o número de colunas que não contenham pivots, da matriz em escada de linhas obtida de A através de operações elementares. Quando um sistema é possível e indeterminado, ao n^o de incógnitas livres desse sistema chama-se **grau de indeterminação** do sistema.

(ii) O número de **incógnitas não livres** de um sistema, é o número de colunas que contenham pivots, da matriz em escada de linhas obtida de A através de operações elementares.

Exemplo 8. O sistema de equações lineares de variáveis reais x, y e z

$$\begin{cases} x + z = 3 \\ x + 2y + 2z = 6 \\ 3y + 3z = 6 \end{cases} \qquad \text{é equivalente a} \qquad \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \\ 0 & 3 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix}.$$

Considere-se então a matriz aumentada e o consequente método de eliminação de Gauss:

$$\begin{bmatrix} 1 & 0 & 1 & | & 3 \\ 1 & 2 & 2 & | & 6 \\ 0 & 3 & 3 & | & 6 \end{bmatrix} \underset{-L_1+L_2 \to L_2}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & | & 3 \\ 0 & 2 & 1 & | & 3 \\ 0 & 3 & 3 & | & 6 \end{bmatrix} \underset{-\frac{3}{2}L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & | & 3 \\ 0 & 2 & 1 & | & 3 \\ 0 & 0 & \frac{3}{2} & | & \frac{3}{2} \end{bmatrix}.$$

Logo,

$$\begin{cases} x + z = 3 \\ 2y + z = 3 \\ \frac{3}{2}z = \frac{3}{2} \end{cases} \Leftrightarrow \begin{cases} x = 2 \\ y = 1 \\ z = 1. \end{cases}$$

Neste exemplo o sistema tem a **solução única** $\{(2,1,1)\}$ e diz-se **possível e determinado**.

Exemplo 9. O sistema de equações lineares de variáveis reais x, y, z e w

$$\begin{cases} 3z - 9w = 6 \\ 5x + 15y - 10z + 40w = -45 \\ x + 3y - z + 5w = -7 \end{cases} \quad \text{é equivalente a} \quad \begin{bmatrix} 0 & 0 & 3 & -9 \\ 5 & 15 & -10 & 40 \\ 1 & 3 & -1 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 6 \\ -45 \\ -7 \end{bmatrix}.$$

Considere-se então a matriz aumentada e o consequente método de eliminação de Gauss:

$$\left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right] \underset{\substack{L_1 \leftrightarrow L_3 \\ \frac{1}{5}L_2 \to L_2}}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{-L_1+L_2 \to L_2}{\longrightarrow}$$

$$\underset{-L_1+L_2 \to L_2}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{3L_2+L_3 \to L_3}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Logo,

$$\begin{cases} x + 3y - z + 5w = -7 \\ -z + 3w = -2 \end{cases} \Leftrightarrow \begin{cases} x = -3y - 2w - 5 \\ z = 3w + 2. \end{cases}$$

As incógnitas y e w são livres e as incógnitas x e z são não livres. A solução geral do sistema é:

$$\left\{ \begin{bmatrix} -3y - 2w - 5 \\ y \\ 3w + 2 \\ w \end{bmatrix} : y, w \in \mathbb{R} \right\}$$

isto é, o conjunto solução é dado por: $\{(-3y - 2w - 5, y, 3w + 2, w) : y, w \in \mathbb{R}\}$. Neste exemplo o sistema tem **um número infinito de soluções** e diz-se **possível e indeterminado** com **grau de indeterminação** 2.

Exemplo 10. Seja $a \in \mathbb{R}$. O sistema de equações lineares de variáveis reais x, y e z

$$\begin{cases} x + 2y + z = 3 \\ x + y - z = 2 \\ x + y + (a^2 - 5)\, z = a \end{cases} \quad \text{é equivalente a} \quad \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & a^2 - 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ a \end{bmatrix}.$$

Considere-se então a matriz aumentada e o consequente método de eliminação de Gauss:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 1 & 1 & -1 & 2 \\ 1 & 1 & a^2 - 5 & a \end{array}\right] \underset{\substack{-L_1+L_2 \to L_2 \\ -L_1+L_3 \to L_3}}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -1 & -2 & -1 \\ 0 & -1 & a^2 - 6 & a - 3 \end{array}\right] \underset{-L_2+L_3 \to L_3}{\longrightarrow}$$

$$\underset{-L_2+L_3 \to L_3}{\longrightarrow} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & -1 & -2 & -1 \\ 0 & 0 & (a-2)(a+2) & a - 2 \end{array}\right].$$

Se $a = 2$, então o sistema é possível e indeterminado:

$$\begin{cases} x + 2y + z = 3 \\ -y - 2z = -1 \end{cases} \Leftrightarrow \begin{cases} x = 3z + 1 \\ y = -2z + 1, \end{cases}$$

a incógnita z é livre, as incógnitas x e y são não livres e a solução geral do sistema é

$$\left\{ \begin{bmatrix} 3z+1 \\ -2z+1 \\ z \end{bmatrix} : z \in \mathbb{R} \right\}$$

isto é, o conjunto solução é dado por: $\{(3z+1, -2z+1, z) : z \in \mathbb{R}\}$.

Assim, se $a = 2$, o sistema tem **um número infinito de soluções** e diz-se **possível e indeterminado** com **grau de indeterminação** 1.

Se $a = -2$, o sistema **não tem solução** e diz-se **impossível**.

Se $a \neq -2$ e $a \neq 2$, o sistema tem a **solução única** $\left\{\left(\frac{a+5}{a+2}, \frac{a}{a+2}, \frac{1}{a+2}\right)\right\}$ e diz-se **possível e determinado.**

Definição 21. (Ver-se-á mais adiante a consistência desta definição.) Seja A uma matriz em escada de linhas. Ao n^o de colunas de A que não contêm pivots chama-se **nulidade** de A e escreve-se $\operatorname{nul} A$. Ao n^o de pivots de A, isto é, ao n^o de linhas não nulas de A, dá-se o nome de **característica** de A e escreve-se $\operatorname{car} A$. Se A fôr a matriz em escada de linhas obtida de C através de operações elementares então diz-se que a **característica** de C é $\operatorname{car} A$, tendo-se $\operatorname{car} C = \operatorname{car} A$ e diz-se que a **nulidade** de C é $\operatorname{nul} A$, tendo-se $\operatorname{nul} C = \operatorname{nul} A$.

Exemplo 11. Considere-se as matrizes do exemplo 2. Pivot de A_1: 4. Pivots de A_2: $1, -5$. Pivots de A_3: $2, -3, -5$. Tem-se: $\operatorname{car} A_1 = 1$, $\operatorname{car} A_2 = 2$ e $\operatorname{car} A_3 = 3$. Além disso: $\operatorname{nul} A_1 = 1$, $\operatorname{nul} A_2 = 2$ e $\operatorname{nul} A_3 = 2$.

Observação 6. Seja $[A \mid B]$ a matriz aumentada associada a um sistema de equações lineares com n incógnitas.

(i) Se $\operatorname{car} A = \operatorname{car} [A \mid B] = n$ então o sistema é **possível e determinado** (tem uma única solução).

(ii) Se $\operatorname{car} A = \operatorname{car} [A \mid B] < n$ então o sistema é **possível e indeterminado** (tem um número infinito de soluções).

(iii) Se $\operatorname{car} A < \operatorname{car} [A \mid B]$ então o sistema é **impossível** (não tem solução).

Observação 7. **(i)** $\operatorname{car} A =$ n^o de linhas não nulas da matriz em escada de linhas obtida de $A =$

$=$ n^o de pivots $=$ n^o de incógnitas não livres.

(ii) $\operatorname{nul} A =$ n^o de incógnitas livres.

Teorema 3. Seja A uma matriz do tipo $m \times n$, isto é, com m linhas e n colunas Então $0 \leq \operatorname{car} A \leq \min\{m, n\}$ e

$$\operatorname{car} A + \operatorname{nul} A = n.$$

Teorema 4. Sejam A uma matriz do tipo $m \times n$ e B uma matriz do tipo $m \times 1$. Se o sistema de equações lineares $AX = B$ tem duas soluções distintas X_0 e X_1 ($X_0 \neq X_1$), então terá um número infinito de soluções.

Dem. Basta verificar que
$$X_\lambda = (1 - \lambda) X_0 + \lambda X_1$$
é solução do sistema $AX = B$, para qualquer $\lambda \in \mathbb{R}$. Além disso, se $\lambda_1 \neq \lambda_2$ então $X_{\lambda_1} \neq X_{\lambda_2}$ uma vez que
$$X_{\lambda_1} - X_{\lambda_2} = (\lambda_2 - \lambda_1)(X_0 - X_1).$$

Teorema 5. Se $A = (a_{ij})_{m \times n}$ é tal que $m < n$, então o sistema linear homogéneo $AX = \mathbf{0}$ tem um número infinito de soluções.

Dem. Como o sistema tem menos equações do que incógnitas ($m < n$), sendo r o n^o de incógnitas não livres, tem-se $n - r$ incógnitas livres as quais podem assumir qualquer valor. Logo, o sistema linear homogéneo $AX = \mathbf{0}$ tem um número infinito de soluções.

Teorema 6. Sejam $A = (a_{ij})_{m \times n}$ e α, β escalares.

(i) Se Y e W são soluções do sistema $AX = \mathbf{0}$, então $Y + W$ também o é.

(ii) Se Y é solução do sistema $AX = \mathbf{0}$, então αY também o é.

(iii) Se Y e W são soluções do sistema $AX = \mathbf{0}$, então $\alpha Y + \beta W$ também o é.

(iv) Sejam Y e W soluções do sistema $AX = B$. Se $\alpha Y + \beta W$ (para quaisquer escalares α, β) também é solução de $AX = B$, então $B = \mathbf{0}$. (Sugestão: basta fazer $\alpha = \beta = 0$.)

Teorema 7. Seja A uma matriz do tipo $m \times n$ e $B \neq \mathbf{0}$ uma matriz do tipo $m \times 1$. Qualquer solução X do sistema $AX = B$ escreve-se na forma $X = X_0 + X_1$ onde X_0 é uma solução particular do sistema $AX = B$ e X_1 é uma solução do sistema linear homogéneo $AX = \mathbf{0}$. Assim:
$$\begin{array}{c}\text{solução geral de} \\ AX = B\end{array} = \begin{array}{c}\text{solução particular de} \\ AX = B\end{array} + \begin{array}{c}\text{solução geral de} \\ AX = \mathbf{0}\end{array}.$$

Dem. Sendo X_0 uma solução particular do sistema $AX = B$ e X_1 uma solução qualquer de $AX = \mathbf{0}$ então
$$A(X_0 + X_1) = AX_0 = B$$
pelo que $X_0 + X_1$ é também uma solução de $AX = B$ e não há solução de $AX = B$ que não seja deste tipo uma vez que, se X' fôr uma solução qualquer de $AX = B$ tem-se
$$AX' = B = AX_0 \Leftrightarrow A(X' - X_0) = \mathbf{0}$$
e assim $X' - X_0 = X_1$ é solução de $AX = \mathbf{0}$ tendo-se $X' = X_0 + X_1$.

Definição 22. Uma matriz A do (tipo $n \times n$) diz-se **invertível** se existir uma matriz B (do tipo $n \times n$) tal que
$$AB = BA = I.$$

À matriz B chama-se **matriz inversa** de A e denota-se por A^{-1}.

Exemplo 12. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ é invertível e $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Observação 8. **(i)** Sendo A^{-1} a matriz inversa de A, então A^{-1} é invertível e a sua inversa é a própria matriz A, isto é,
$$\left(A^{-1}\right)^{-1} = A.$$

(ii) A matriz nula não é invertível. No entanto, a matriz identidade I é invertível tendo-se
$$I^{-1} = I.$$

(iii) Se uma matriz quadrada tiver uma linha ou uma coluna nula então não é invertível.

Teorema 8. A inversa de uma matriz invertível é única.

Dem. Sejam B e C as inversas de A. Então,
$$B = BI = B\,(AC) = (BA)\,C = IC = C.$$

Definição 23. **(i)** Uma matriz A diz-se **simétrica** se $A = A^T$, isto é, se
$$a_{ij} = a_{ji},$$
para $i, j = 1, ..., n$. Diz-se que A é **anti-simétrica** se $A = -A^T$, isto é, se
$$a_{ij} = -a_{ji},$$
para $i, j = 1, ..., n$.

Exemplo 13. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ é uma matriz simétrica. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Teorema 9. **(i)** Se $A = (a_{ij})_{n\times n}$ e $B = (b_{ij})_{n\times n}$ são duas matrizes invertíveis, então AB é invertível e
$$(AB)^{-1} = B^{-1}A^{-1}.$$

(ii) Sendo α um escalar não nulo e A uma matriz invertível então αA é invertível e

$$(\alpha A)^{-1} = \frac{1}{\alpha} A^{-1}.$$

(iii) Seja $m \in \mathbb{N}$. Se $A = (a_{ij})_{n\times n}$ é uma matriz invertível, então A^m é invertível e

$$(A^m)^{-1} = \left(A^{-1}\right)^m$$

e escreve-se

$$A^{-m} = (A^m)^{-1}.$$

(iv) Seja $A = (a_{ij})_{n\times n}$ uma matriz. Se existir $l \in \mathbb{N}$ tal que $A^l = \mathbf{0}$ então A não é invertível.

(v) Sejam A e B matrizes com A invertível tais que $AB = \mathbf{0}$. Então $B = \mathbf{0}$.

(vi) Sejam A e B matrizes com B invertível tais que $AB = \mathbf{0}$. Então $A = \mathbf{0}$.

(vii) Sejam A, B e C matrizes com A invertível tais que $AB = AC$. Então $B = C$.

(viii) Sejam A, B e C matrizes com B invertível tais que $AB = CB$. Então $A = C$.

(ix) $A = (a_{ij})_{n\times n}$ é uma matriz invertível se e só se A^T é invertível e

$$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T.$$

(x) Se $A = (a_{ij})_{n\times n}$ é uma matriz simétrica invertível, então A^{-1} é simétrica.

(xi) Se A e B são duas matrizes simétricas então AB é uma matriz simétrica se e só se A e B comutarem.

Teorema 10. Seja A uma matriz do tipo $n \times n$.

(i) O sistema $AX = B$ tem solução única se e só se A fôr invertível. Neste caso a solução geral é $X = A^{-1}B$.

(ii) O sistema homogéneo $AX = \mathbf{0}$ tem solução não trivial se e só se A fôr não invertível.

Teorema 11. **(i)** Sejam A e B duas matrizes do tipo $n \times n$. Se AB é invertível, então A e B são invertíveis.

(ii) Se A é uma matriz do tipo $n \times n$ tal que $AB = I$ então $BA = I$ e $B = A^{-1}$.

Dem. **(i)** Considere o sistema $(AB)\,X = \mathbf{0}$. Se B não fosse invertível, então existiria $X \neq \mathbf{0}$ tal que $BX = \mathbf{0}$. Logo, $X \neq \mathbf{0}$ seria solução não trivial de $ABX = \mathbf{0}$, o que contraria o teorema anterior uma vez que por hipótese AB é invertível. Assim, B é invertível. Finalmente, A é invertível por ser o produto de duas matrizes invertíveis: $A = (AB)\, B^{-1}$.

(ii) Atendendo à alínea anterior, B é invertível. Logo B^{-1} também é invertível e

$$A = AI = A\left(BB^{-1}\right) = (AB)\, B^{-1} = IB^{-1} = B^{-1},$$

isto é, A é invertível e $A^{-1} = \left(B^{-1}\right)^{-1} = B$.

Teorema 12. (**Como inverter matrizes invertíveis do tipo $n \times n$**). Seja A uma matriz do tipo $n \times n$ e consideremos a equação $AX = B$. **Se A fôr invertível** temos

$$AX = B \Leftrightarrow X = A^{-1}B,$$

isto é,

$$AX = IB \Leftrightarrow IX = A^{-1}B.$$

Assim, para determinar a inversa de A, iremos transformar a matriz aumentada $[A \mid I]$ na matriz $[I \mid A^{-1}]$, por meio de operações elementares aplicadas às linhas de $[A \mid I]$:

$$[A \mid I] \underset{\cdots}{\longrightarrow} \left[I \mid A^{-1}\right]$$

Este método tem o nome de **método de eliminação de Gauss-Jordan** e consistirá na continuação do método de eliminação de Gauss agora aplicado a [matriz triangular superior | $*$], efectuando-se as eliminações de baixo para cima de modo a obter-se $[I \mid A^{-1}]$.

Exemplo 14. Vejamos que $\begin{bmatrix} -2 & 1 \\ -1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$. Tem-se

$$\left[\begin{array}{cc|cc} -2 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{array}\right] \underset{-\frac{1}{2}L_1+L_2 \to L_2}{\longrightarrow} \left[\begin{array}{cc|cc} -2 & 1 & 1 & 0 \\ 0 & \frac{3}{2} & -\frac{1}{2} & 1 \end{array}\right] \underset{-\frac{2}{3}L_2+L_1 \to L_1}{\longrightarrow}$$

$$\underset{-\frac{2}{3}L_2+L_1 \to L_1}{\longrightarrow} \left[\begin{array}{cc|cc} -2 & 0 & \frac{4}{3} & -\frac{2}{3} \\ 0 & \frac{3}{2} & -\frac{1}{2} & 1 \end{array}\right] \underset{\substack{\frac{2}{3}L_2 \to L_2 \\ -\frac{1}{2}L_1 \to L_1}}{\longrightarrow} \left[\begin{array}{cc|cc} 1 & 0 & -\frac{2}{3} & \frac{1}{3} \\ 0 & 1 & -\frac{1}{3} & \frac{2}{3} \end{array}\right].$$

Isto é

$$\begin{bmatrix} -2 & 1 \\ -1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

De facto

$$\begin{bmatrix} -2 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} -2 & 1 \\ -1 & 2 \end{bmatrix} = I$$

Exemplo 15. **(i)** Seja $A = \begin{bmatrix} 0 & -1 & 1 \\ -1 & \frac{5}{4} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & 0 \end{bmatrix}$. Tem-se

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 0 & -1 & 1 & 1 & 0 & 0 \\ -1 & \frac{5}{4} & -\frac{1}{2} & 0 & 1 & 0 \\ 1 & -\frac{1}{2} & 0 & 0 & 0 & 1 \end{array}\right] \underset{\cdots}{\longrightarrow} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 2 & 3 \\ 0 & 1 & 0 & 2 & 4 & 4 \\ 0 & 0 & 1 & 3 & 4 & 4 \end{array}\right].$$

Logo, $\begin{bmatrix} 0 & -1 & 1 \\ -1 & \frac{5}{4} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \\ 3 & 4 & 4 \end{bmatrix}$. Verifique(!) que: $AA^{-1} = I$.

(ii) Seja $A = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$. Tem-se $[A \mid I] \underset{\cdots}{\longrightarrow} \left[\begin{array}{ccc|ccc} 3 & 2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right]$. Logo, A não é invertível.

(iii) Sejam $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $B = \begin{bmatrix} -4 & 0 \\ 0 & 8 \end{bmatrix}$ $C = \begin{bmatrix} 0 & \frac{1}{8} \\ -\frac{1}{4} & 0 \end{bmatrix}$. Determine-se X tal que

$$A\left(I - 2X^T\right)^{-1} B^{-1} = C.$$

Tem-se $A\left(I - 2X^T\right)^{-1} B^{-1} = C \Leftrightarrow \left(I - 2X^T\right)^{-1} = A^{-1}CB \Leftrightarrow I - 2X^T = \left(A^{-1}CB\right)^{-1} \Leftrightarrow$

$$\Leftrightarrow X^T = \frac{1}{2}\left(I - B^{-1}C^{-1}A\right) \Leftrightarrow X = \frac{1}{2}\left(I - A^T\left(C^T\right)^{-1}\left(B^T\right)^{-1}\right) \Leftrightarrow$$

$$\Leftrightarrow X = \frac{1}{2}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 0 & -\frac{1}{4} \\ \frac{1}{8} & 0 \end{bmatrix}^{-1}\begin{bmatrix} -4 & 0 \\ 0 & 8 \end{bmatrix}^{-1}\right) \Leftrightarrow X = \begin{bmatrix} -1 & -\frac{1}{2} \\ -2 & -\frac{1}{2} \end{bmatrix}.$$

Teorema 13. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e considere o sistema de equações lineares $AX = B$.

(i) Existência de solução: Se $m \leq n$ então o sistema $AX = B$ tem pelo menos uma solução X para cada $B \in \mathbb{R}^m$ se e só se

$$\operatorname{car} A = m.$$

(ii) Unicidade de solução: Se $m \geq n$ então o sistema $AX = B$ tem no máximo uma solução X para cada $B \in \mathbb{R}^m$ se e só se

$$\operatorname{car} A = n,$$

isto é, se e só se $\operatorname{nul} A = 0$.

(iii) Existência e unicidade de solução: Se $m = n$ então:

$$A \text{ é invertível} \Leftrightarrow \operatorname{car} A = n \Leftrightarrow$$

$$\Leftrightarrow \text{para todo o } B \text{ o sistema } AX = B \text{ tem uma única solução } (X = A^{-1}B),$$

isto é,

$$A \text{ não é invertível} \Leftrightarrow \operatorname{car} A < n \Leftrightarrow$$

$$\Leftrightarrow \text{existe pelo menos um } B \text{ para o qual o sistema } AX = B \text{ não tem solução.}$$

Espaços lineares (ou Espaços vectoriais)

Definição 24. Um conjunto não vazio V é um **espaço linear** (real) se existirem duas operações associadas a V, uma soma de elementos de V e um produto de escalares (números reais) por elementos de V, com as seguintes propriedades:

(a) (Fecho da soma). Para quaisquer $u, v \in V$

$$u + v \in V.$$

(b) (Fecho do produto por escalares). Para quaisquer $\alpha \in \mathbb{R}$ e $u \in V$

$$\alpha u \in V.$$

(c) (Comutatividade da soma). Para quaisquer $u, v \in V$,

$$u + v = v + u.$$

(d) (Associatividade da soma). Para quaisquer $u, v, w \in V$,

$$u + (v + w) = (u + v) + w.$$

(e) (Elemento neutro da soma). Existe um elemento de V designado por $\mathbf{0}$ tal que, para qualquer $u \in V$,

$$u + \mathbf{0} = u.$$

(f) (Simétrico). Para cada (qualquer) $u \in V$ existe $v \in V$ tal que

$$u + v = \mathbf{0}.$$

A v chama-se o **simétrico** de u e denota-se por $-u$.

(g) (Associatividade do produto por escalares). Para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u \in V$,

$$\alpha (\beta u) = (\alpha \beta) u.$$

(h) (Distributividade em relação à soma de vectores). Para quaisquer $\alpha \in \mathbb{R}$ e $u, v \in V$,

$$\alpha (u + v) = \alpha u + \alpha v.$$

(i) (Distributividade em relação à soma de escalares). Para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u \in V$,

$$(\alpha + \beta) u = \alpha u + \beta u.$$

(j) Para qualquer $u \in V$,

$$1u = u.$$

Definição 25. Aos elementos de um espaço linear (vectorial) V chamaremos vectores.

Exemplo 16. Exemplos de espaços lineares. Seja $\alpha \in \mathbb{R}$.

(i) $\mathbb{R}^n = \{(x_1, ..., x_n) : x_1, ..., x_n \in \mathbb{R}\}$, com as operações usuais:

$$(u_1, ..., u_n) + (v_1, ..., v_n) = (u_1 + v_1, ..., u_n + v_n),$$

$$\alpha(u_1, ..., u_n) = (\alpha u_1, ..., \alpha u_n).$$

(ii) $\mathcal{M}_{m\times n}(\mathbb{R})$ (conjunto de todas as matrizes reais do tipo $m \times n$), com as operações (usuais): $A + B$ e αA.

(iii) Seja $n \in \mathbb{N}$ fixo. O conjunto $\mathcal{P}_n = \{a_0 + a_1t + ... + a_nt^n : a_0, a_1, ..., a_n \in \mathbb{R}\}$ de todos os polinómios reais de variável real e de grau menor ou igual a n, com as operações usuais.

$$(a_0 + a_1t + ... + a_nt^n) + (b_0 + b_1t + ... + b_nt^n) = a_0 + b_0 + (a_1 + b_1)\, t + ... + (a_n + b_n)\, t^n$$

$$\alpha\,(a_0 + a_1t + ... + a_nt^n) = \alpha a_0 + (\alpha a_1)\, t + ... + (\alpha a_n)\, t^n.$$

(iv) O conjunto $\mathcal{P} = \{a_0 + a_1t + ... + a_st^s : a_0, a_1, ..., a_s \in \mathbb{R} \text{ e } s \in \mathbb{N}_0\}$ de todos os polinómios reais de variável real, com as operações usuais.

(v) O conjunto de todas as funções reais de variável real definidas num conjunto $S \subset \mathbb{R}$, com as operações usuais:

$$(f + g)(x) = f(x) + g(x),$$
$$(\alpha f)(x) = \alpha f(x).$$

Observação 9. Existem espaços lineares com operações não usuais:

(i) O conjunto dos números reais $\mathbb{R}$, com a soma definida por

$$u \boxplus v = u + v + 1,$$

e o produto por escalares definido por

$$\alpha \cdot u = \alpha u + \alpha - 1,$$

é um espaço linear. (Neste caso o elemento neutro é -1.)

(ii) O conjunto dos números reais maiores do que zero, com a soma definida por

$$u \boxplus v = uv,$$

e o produto por escalares definido por

$$\alpha \cdot u = u^\alpha,$$

é um espaço linear. (Neste caso o elemento neutro é 1.)

Observação 10. Alterações nos conjuntos considerados anteriormente podem resultar em conjuntos que não são espaços lineares.

(i) O conjunto $\{(x, y) \in \mathbb{R}^2 : x \geq 0 \text{ e } y \geq 0\}$, com as operações usuais, não é um espaço linear. Por exemplo, os simétricos não estão no conjunto.

(ii) O conjunto $V = \{a_0 + a_1 t + ... + a_n t^n : a_0, a_1, ..., a_n \in \mathbb{R} \text{ e } a_n \neq 0\}$ de todos os polinómios reais de grau igual a n, com as operações usuais, não é um espaço linear. Por exemplo, para $n > 1$:

$$t^n, -t^n + t \in V, \quad \text{mas} \quad t^n + (-t^n + t) = t \notin V.$$

(iii) O conjunto $U = \{f : \mathbb{R} \longrightarrow \mathbb{R} \text{ tais que } f(1) = 2\}$, com as operações usuais, não é um espaço linear. Por exemplo, se $f_1, f_2 \in U$,

$$(f_1 + f_2)(1) = f_1(1) + f_2(1) = 2 + 2 = 4 \neq 2.$$

Logo, $f_1 + f_2 \notin U$.

Definição 26. Seja V um espaço linear. Diz-se que S é um **subespaço** de V se S é um subconjunto de V e se S, com as operações de V, fôr um espaço linear.

Observação 11. No entanto, para mostrar que um certo conjunto $S \subset V$ é um subespaço do espaço linear V, não será necessário verificar as 10 propriedades da definição de espaço linear, como se pode ver no seguinte teorema.

Teorema 14. Um subconjunto não vazio S de um espaço linear V é um subespaço de V se e só se as seguintes condições (i) e (ii) forem satisfeitas.

(i) Para quaisquer $u, v \in S$ tem-se $u + v \in S$.

(ii) Para quaisquer $\alpha \in \mathbb{R}$ e $u \in S$ tem-se $\alpha u \in S$.

Exemplo 17. Exemplos de subespaços:

(i) Os únicos subespaços do espaço linear $\mathbb{R}$, com as operações usuais, são $\{0\}$ e $\mathbb{R}$.

(ii) Os subespaços do espaço linear $\mathbb{R}^3$, com as operações usuais, são: $\{(0, 0, 0)\}$, $\mathbb{R}^3$, todas as rectas que passam pela origem e todos os planos que passam pela origem.

(iii) O conjunto de todas as matrizes (reais) triangulares superiores (do tipo $n \times n$) é um subespaço do espaço linear $\mathcal{M}_{n \times n}(\mathbb{R})$, com as operações usuais.

(iv) O conjunto de todas as funções reais definidas e contínuas em $I \subset \mathbb{R}$ (I é um intervalo) é um subespaço do espaço linear de todas as funções $f : I \to \mathbb{R}$, com as operações usuais.

Definição 27. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. O conjunto

$$\mathcal{C}(A) = \{b \in \mathbb{R}^m : Au = b \text{ tem pelo menos uma solução } u\}$$

é um subespaço do espaço linear $\mathbb{R}^m$, com as operações usuais, ao qual se dá o nome de **espaço das colunas** de A.

Definição 28. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. O conjunto

$$\mathcal{N}(A) = \{u \in \mathbb{R}^n : Au = \mathbf{0}\}$$

é um subespaço do espaço linear $\mathbb{R}^n$, com as operações usuais, ao qual se dá o nome de **núcleo** de A.

Teorema 15 . Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$.

$$A \text{ invertível } \Leftrightarrow \mathcal{N}(A) = \{\mathbf{0}\}$$

Definição 29. Seja S um subconjunto não vazio de um espaço linear V. Diz-se que um vector u é **combinação linear** finita dos elementos de S, se existir um n^o finito de elementos de S, $u_1, ..., u_k$, e de escalares $\lambda_1, ..., \lambda_k$ tais que

$$u = \lambda_1 u_1 + ... + \lambda_k u_k = \sum_{i=1}^{k} \lambda_i u_i.$$

Seja

$$L(S) = \{\lambda_1 u_1 + ... + \lambda_k u_k : \lambda_1, ..., \lambda_k \in \mathbb{R}\},$$

(no caso do corpo dos escalares ser $\mathbb{R}$) isto é, seja $L(S)$ o conjunto de todas as combinações lineares finitas de elementos de S. O conjunto $L(S)$ é (verifique!) um subespaço de V. A $L(S)$ chama-se a **expansão linear** de S ou **subespaço de** V **gerado** por S e diz-se que S **gera** $L(S)$ ou ainda que S é um **conjunto gerador** do espaço linear $L(S)$. Se S é o conjunto vazio $\varnothing$, escreve-se $L(\varnothing) = \{\mathbf{0}\}$.

Teorema 16. (i) Seja S um subconjunto não vazio de um espaço linear V. A expansão linear $L(S)$ de S é o menor subespaço de V que contém S.

(ii) Sejam S e T dois subconjuntos não vazios de um espaço linear V, com $S \subset T$. Se $L(S) = V$ então $L(T) = V$.

Definição 30. Seja A uma matriz (real) do tipo $m \times n$. Ao subespaço linear de $\mathbb{R}^n$ gerado pelas linhas de A dá-se o nome de **espaço das linhas** de A e designa-se por $\mathcal{L}(A)$.

Exemplo 18. (i) O espaço linear $\mathbb{R}^2$ é gerado por qualquer dos seguintes conjuntos de vectores:

$$\{(1,0),(0,1)\}, \quad \{(1,2),(-1,11)\} \quad \text{e} \quad \{(23,8),(6,14)\}.$$

(ii) O subespaço $\{(x,y) \in \mathbb{R}^2 : y = 2x\}$ do espaço linear $\mathbb{R}^2$ é gerado por qualquer dos seguintes conjuntos de vectores:

$$\{(1,2)\}, \quad \{(-2,-4)\} \quad \text{e} \quad \{(77,154)\}.$$

(iii) O espaço linear $\mathcal{P}_n$ de todos os polinómios reais de variável real e de grau menor ou igual a n, é gerado por qualquer dos seguintes conjuntos de vectores:

$$\{1, t, t^2, ..., t^n\}, \quad \{1, 1+t, (1+t)^2, ..., (1+t)^n\} \quad \text{e} \quad \{1, \frac{t}{1!}, \frac{t^2}{2!}, ..., \frac{t^n}{n!}\}.$$

(iv) O espaço linear $\mathcal{P}$ de todos os polinómios reais de variável real, é gerado pelo conjunto infinito de vectores:

$$\{1, t, t^2, ...\}.$$

(v) Seja U o espaço linear de todas as funções reais com primeira derivada contínua em $\mathbb{R}$ (isto é, pertencentes a $C^1(\mathbb{R})$) e tais que $f'(x) = af(x)$ (em $\mathbb{R}$) com $a \in \mathbb{R}$. Então U é gerado pela função $g(x) = e^{ax}$, tendo-se $U = L(\{g\})$.

(vi) Seja A uma matriz (real) do tipo $m \times n$. O espaço das colunas de A,

$$\mathcal{C}(A) = \{b \in \mathbb{R}^m : Au = b \text{ tem pelo menos uma solução } u\},$$

é o subespaço (do espaço linear $\mathbb{R}^m$) gerado pelas colunas de A, uma vez que:

$$\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = u_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + ... + u_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

(vii) $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -3 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{bmatrix}$, $C = \begin{bmatrix} -1 & 2 \\ 2 & -4 \\ -2 & 4 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$.

$$\mathcal{C}(A) = \{(0,0)\}, \qquad \mathcal{N}(A) = \mathbb{R}^3, \qquad \mathcal{L}(A) = \{(0,0,0)\}.$$

$$\mathcal{C}(B) = L(\{(1,0,0),(1,7,0)\}), \quad \mathcal{N}(B) = L(\{(3,1,0)\}), \quad \mathcal{L}(B) = L(\{(1,-3,1),(0,0,7)\}).$$

$$\mathcal{C}(C) = L(\{(-1,2,-2)\}), \qquad \mathcal{N}(C) = L(\{(2,1)\}), \qquad \mathcal{L}(C) = L(\{(-1,2)\}).$$

$$\mathcal{C}(D) = L(\{(2,0),(0,-1)\}), \ \mathcal{N}(D) = \{(0,0)\}, \qquad \mathcal{L}(D) = L(\{(2,0),(0,-1)\}).$$

(viii) Seja $U = \{A \in \mathcal{M}_{3\times 2}(\mathbb{R}) : a_{12} = a_{21} = a_{32} = 0 \text{ e } a_{11} + 2a_{31} = 0\}$. Tem-se, para $A \in U$,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} -2a_{31} & 0 \\ 0 & a_{22} \\ a_{31} & 0 \end{bmatrix} = a_{31} \begin{bmatrix} -2 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

com $a_{31}, a_{22} \in \mathbb{R}$. Logo,

$$U = L\left(\left\{\begin{bmatrix} -2 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\right\}\right).$$

(ix) Seja $U = \{p(t) = a_0 + a_1 t + a_2 t^2 \in \mathcal{P}_2 : p(1) = p(0)\}$. Tem-se, para $p(t) \in U$,

$$p(1) = p(0) \Leftrightarrow a_0 + a_1 + a_2 = a_0 \Leftrightarrow a_1 + a_2 = 0 \Leftrightarrow a_1 = -a_2.$$

Logo, $p(t) = a_0 - a_2 t + a_2 t^2 = a_0 1 + a_2 \left(-t + t^2\right)$, com $a_0, a_2 \in \mathbb{R}$. Assim,

$$U = L\left(\left\{1, -t + t^2\right\}\right).$$

Teorema 17. Se U e V subespaços do espaço linear W, então $U \cup V$ é subespaço de W se e só se $U \subset V$ ou $V \subset U$.

Teorema 18. Se U e V são subespaços do espaço linear W, então:

(i) O conjunto $U \cap V$ é um subespaço linear de W.

(ii) O conjunto $U + V = \{u + v : u \in U \quad \text{e} \quad v \in V\}$ é um subespaço de W. É o menor subespaço de W que contém $U \cup V$. Tem-se $U + V = L(U \cup V)$. O conjunto $U \cup V$ em geral não é um subespaço.

Observação 12. **(i)** U é um subespaço de $\mathbb{R}^n$ se e só se existir uma matriz A tal que

$$U = \mathcal{N}(A).$$

(ii) Sejam U_1 e U_2 subespaços de $\mathbb{R}^n$. Se $U_1 = L(S_1)$ e $U_2 = L(S_2)$ então

$$U_1 + U_2 = L(S_1 \cup S_2).$$

Se $U_1 = \mathcal{N}(A)$ e $U_2 = \mathcal{N}(B)$ então

$$U_1 \cap U_2 = \mathcal{N}\begin{pmatrix} A \\ B \end{pmatrix}.$$

(iii) U é um subespaço de $\mathcal{P}_n = \{a_0 + a_1 t + ... + a_n t^n : a_0, a_1, ..., a_n \in \mathbb{R}\}$ se e só se existir uma matriz A tal que

$$U = \{a_0 + a_1 t + ... + a_n t^n : (a_0, a_1, ..., a_n) \in \mathcal{N}(A)\}.$$

(iv) Sejam U_1 e U_2 subespaços de $\mathcal{P}_n$. Se $U_1 = L(S_1)$ e $U_2 = L(S_2)$ então

$$U_1 + U_2 = L(S_1 \cup S_2).$$

Se

$$U_1 = \{a_0 + a_1 t + ... + a_n t^n : (a_0, a_1, ..., a_n) \in \mathcal{N}(A)\}$$

e

$$U_2 = \{a_0 + a_1 t + ... + a_n t^n : (a_0, a_1, ..., a_n) \in \mathcal{N}(B)\}$$

então

$$U_1 \cap U_2 = \left\{a_0 + a_1 t + ... + a_n t^n : (a_0, a_1, ..., a_n) \in \mathcal{N}\begin{pmatrix} A \\ B \end{pmatrix}\right\}.$$

(v) U é um subespaço de $\mathcal{M}_{m\times n}(\mathbb{R})$ se e só se existir uma matriz B tal que

$$U = \{A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{R}) : (a_{11}, ..., a_{m1}, ..., a_{1n}, ..., a_{mn}) \in \mathcal{N}(B)\}$$

(vi) Sejam U_1 e U_2 subespaços de $\mathcal{M}_{m\times n}(\mathbb{R})$. Se $U_1 = L(S_1)$ e $U_2 = L(S_2)$ então

$$U_1 + U_2 = L(S_1 \cup S_2).$$

Se

$$U_1 = \{A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{R}) : (a_{11}, ..., a_{m1}, ..., a_{1n}, ..., a_{mn}) \in \mathcal{N}(B)\}$$

e

$$U_2 = \{A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{R}) : (a_{11}, ..., a_{m1}, ..., a_{1n}, ..., a_{mn}) \in \mathcal{N}(C)\}$$

então

$$U_1 \cap U_2 = \left\{A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{R}) : (a_{11}, ..., a_{m1}, ..., a_{1n}, ..., a_{mn}) \in \mathcal{N}\begin{pmatrix} B \\ C \end{pmatrix}\right\}.$$

Exemplo 19. Em $\mathbb{R}^3$, considere os subespaços:

$$U = L(\{(1,-1,1),(1,2,2)\}) \quad \text{e} \quad V = L(\{(2,1,1),(-1,1,3)\}).$$

Seja

$$(x,y,z) \in U = L(\{(1,-1,1),(1,2,2)\}).$$

Assim, existem escalares $\alpha, \beta \in \mathbb{R}$ tais que

$$(x,y,z) = \alpha(1,-1,1) + \beta(1,2,2).$$

Logo, o sistema seguinte é possível

$$\left[\begin{array}{cc|c} 1 & 1 & x \\ -1 & 2 & y \\ 1 & 2 & z \end{array}\right].$$

Atendendo a que

$$\left[\begin{array}{cc|c} 1 & 1 & x \\ -1 & 2 & y \\ 1 & 2 & z \end{array}\right] \underset{\substack{L_1+L_2\to L_2 \\ -L_1+L_3\to L_3}}{\longrightarrow} \left[\begin{array}{cc|c} 1 & 1 & x \\ 0 & 3 & x+y \\ 0 & 1 & z-x \end{array}\right] \underset{-\frac{1}{3}L_2+L_3\to L_3}{\longrightarrow} \left[\begin{array}{cc|c} 1 & 1 & x \\ 0 & 3 & x+y \\ 0 & 0 & z-\frac{4}{3}x-\frac{1}{3}y \end{array}\right]$$

logo

$$(x,y,z)\in U \Leftrightarrow z-\frac{4}{3}x-\frac{1}{3}y=0 \Leftrightarrow 4x+y-3z=0.$$

Ou seja:

$$U=\{(x,y,z)\in\mathbb{R}^3 : 4x+y-3z=0\}=\mathcal{N}\left(\left[\begin{array}{ccc} 4 & 1 & -3 \end{array}\right]\right).$$

Seja

$$(x,y,z)\in V=L\left(\{(2,1,1),(-1,1,3)\}\right).$$

Existem escalares $\alpha,\beta\in\mathbb{R}$ tais que

$$(x,y,z)=\alpha(2,1,1)+\beta(-1,1,3).$$

Logo, o sistema seguinte é possível

$$\left[\begin{array}{cccc} 2 & -1 & | & x \\ 1 & 1 & | & y \\ 1 & 3 & | & z \end{array}\right].$$

Atendendo a que

$$\left[\begin{array}{cccc} 2 & -1 & | & x \\ 1 & 1 & | & y \\ 1 & 3 & | & z \end{array}\right] \underset{\substack{-\frac{1}{2}L_1+L_2\to L_2 \\ -\frac{1}{2}L_1+L_3\to L_3}}{\longrightarrow} \left[\begin{array}{cccc} 2 & -1 & | & x \\ 0 & 3/2 & | & y-\frac{x}{2} \\ 0 & 7/2 & | & z-\frac{x}{2} \end{array}\right] \underset{-\frac{7}{3}L_2+L_3\to L_3}{\longrightarrow} \left[\begin{array}{cccc} 2 & -1 & | & x \\ 0 & 3/2 & | & y-\frac{x}{2} \\ 0 & 0 & | & \frac{2}{3}x-\frac{7}{3}y+z \end{array}\right]$$

logo

$$(x,y,z)\in V \Leftrightarrow z-\frac{7}{3}y+\frac{2}{3}x=0 \Leftrightarrow 2x-7y+3z=0.$$

Ou seja:

$$V=\{(x,y,z)\in\mathbb{R}^3 : 2x-7y+3z=0\}=\mathcal{N}\left(\left[\begin{array}{ccc} 2 & -7 & 3 \end{array}\right]\right).$$

Logo,

$$\begin{aligned} U\cap V &= \mathcal{N}\left(\left[\begin{array}{ccc} 4 & 1 & -3 \\ 2 & -7 & 3 \end{array}\right]\right)=\mathcal{N}\left(\left[\begin{array}{ccc} 4 & 1 & -3 \\ 0 & -5 & 3 \end{array}\right]\right)= \\ &= \{(3y,3y,5y) : \mu\in\mathbb{R}\}=\{y(3,3,5) : y\in\mathbb{R}\}=L\left(\{(3,3,5)\}\right). \end{aligned}$$

(iii) Seja U o subespaço de $\mathcal{M}_{n\times n}(\mathbb{R})$ das matrizes triangulares superiores e seja V o subespaço de $\mathcal{M}_{n\times n}(\mathbb{R})$ das matrizes triangulares inferiores. Então

$$U+V=\mathcal{M}_{n\times n}(\mathbb{R}) \quad \text{e} \quad U\cap V= \text{ subespaço das matrizes diagonais.}$$

(iv) Sejam $U=L(\{(1,0)\})$ e $V=L(\{(0,1)\})$ subespaços de $\mathbb{R}^2$. O conjunto

$$U\cup V=\{(x,y)\in\mathbb{R}^2 : x=0\vee y=0\}$$

não é um espaço linear pois $\underbrace{(1,0)}_{\in U}+\underbrace{(0,1)}_{\in V}=(1,1)\notin U\cup V$. No entanto, tem-se $U+V=\mathbb{R}^2$.

Observação 13. Vejamos que se tem:

$$L\left(\{(1,-4,0),(0,3,1)\}\right) = L\left(\{(1,2,2),(1,-1,1)\}\right).$$

Como

$$(1,-4,0) = -(1,2,2) + 2(1,-1,1) \quad \text{e} \quad (0,3,1) = (1,2,2) - (1,-1,1)$$

logo

$$L\left(\{(1,-4,0),(0,3,1)\}\right) \subset L\left(\{(1,2,2),(1,-1,1)\}\right).$$

Como

$$(1,2,2) = (1,-4,0) + 2(0,3,1) \quad \text{e} \quad (1,-1,1) = (1,-4,0) + (0,3,1)$$

logo

$$L\left(\{(1,2,2),(1,-1,1)\}\right) \subset L\left(\{(1,-4,0),(0,3,1)\}\right).$$

Assim:

$$L\left(\{(1,-4,0),(0,3,1)\}\right) = L\left(\{(1,2,2),(1,-1,1)\}\right).$$

De facto, o que se mostrou foi o seguinte:

$$\begin{bmatrix} 1 & 0 \\ -4 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 0 \\ -4 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ 2 & 1 \end{bmatrix}$$

em que

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}^{-1}.$$

Definição 31. **(i)** Sejam W_1 e W_2 subespaços de um espaço linear V. Diz-se que V é a **soma directa** dos espaços W_1 e W_2 e escreve-se

$$V = W_1 \oplus W_2$$

se

$$V = W_1 + W_2 \qquad \text{e} \qquad W_1 \cap W_2 = \{\mathbf{0}\}.$$

(ii) Sejam $W_1, ..., W_k$ subespaços de um espaço linear V. Diz-se que V é a **soma directa** dos espaços $W_1, ..., W_k$ e escreve-se

$$V = W_1 \oplus ... \oplus W_k$$

se

$$V = W_1 + ... + W_k \qquad \text{e} \qquad W_r \cap \sum_{\substack{i=1 \\ i \neq r}}^{k} W_i = \{\mathbf{0}\}, \qquad \text{para todo o } r = 1, ..., k.$$

Teorema 19. **(i)** Sejam W_1 e W_2 subespaços de um espaço linear V tais que $V = W_1 \oplus W_2$. Então todo o vector $v \in V$ pode ser escrito de modo único na forma

$$v = w_1 + w_2$$

com $w_1 \in W_1$ e $w_2 \in W_2$.

(ii) Sejam $W_1, ..., W_k$ subespaços de um espaço linear V tais que $V = W_1 \oplus ... \oplus W_k$. Então todo o vector $v \in V$ pode ser escrito de modo único na forma

$$v = w_1 + ... + w_k$$

com $w_i \in W_i$, para todo o $i = 1, ..., k$.

Teorema 20. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Tem-se

$$\mathcal{C}(A) = \mathcal{L}(A^T) \qquad \text{e} \qquad \mathcal{L}(A) \cap \mathcal{N}(A) = \{\mathbf{0}\}.$$

Dem. Vejamos que

$$\mathcal{N}(A) \cap \mathcal{C}\left(A^T\right) = \{\mathbf{0}\}.$$

Seja

$$y \in \mathcal{N}(A) \cap \mathcal{C}\left(A^T\right).$$

Então

$$Ay = \mathbf{0} \quad \text{e} \quad \text{existe } x \text{ tal que } y = A^T x.$$

Logo

$$y^T = x^T A e y^T y = \left(x^T A\right) y = x^T (Ay) = x^T \mathbf{0} = 0.$$

Isto é

$$\sum_{i=1}^{n} y_i^2 = y^T y = 0$$

ou seja

$$y = (y_1, ..., y_n) = (0, ..., 0) = \mathbf{0}.$$

Logo

$$\mathcal{N}(A) \cap \mathcal{L}(A) = \mathcal{N}(A) \cap \mathcal{C}\left(A^T\right) = \{\mathbf{0}\}.$$

Observação 14. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Se A' fôr a matriz em escada que se obtem de A por aplicação do método de eliminação de Gauss, tem-se

$$\mathcal{C}(A) \neq \mathcal{C}(A').$$

Teorema 21. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. O espaço das linhas $\mathcal{L}(A)$ e o núcleo $\mathcal{N}(A)$ mantêm-se invariantes por aplicação do método de eliminação de Gauss. Isto é, sendo A' a matriz em escada que se obtem de A por aplicação desse método, tem-se

$$\mathcal{L}(A) = \mathcal{L}(A') \qquad \text{e} \qquad \mathcal{N}(A) = \mathcal{N}(A').$$

Independência linear

Definição 32. **(i)** Seja V um espaço linear. Seja

$$S = \{v_1, ..., v_k\} \subset V.$$

Diz-se que o conjunto S é **linearmente dependente** se e só se algum dos vectores de S se escrever como combinação linear dos restantes, isto é, se e só se existir algum $i \in \{1, ..., k\}$ e escalares $\lambda_1, ..., \lambda_{i-1}, \lambda_{i+1}, ..., \lambda_k \in \mathbb{R}$ tais que

$$v_i = \lambda_1 v_1 + ... + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + ... + \lambda_k v_k.$$

(ii) Seja V um espaço linear. Seja

$$S = \{v_1, ..., v_k\} \subset V.$$

Diz-se que o conjunto S é **linearmente independente** se e só se nenhum dos vectores de S se puder escrever como combinação linear dos restantes, isto é, se e só a única solução do sistema homogéneo

$$\lambda_1 v_1 + ... + \lambda_k v_k = \mathbf{0}$$

fôr a solução trivial, ou seja, $\lambda_1 = ... = \lambda_k = 0$. No caso em que $V = \mathbb{R}^n$, sendo A a matriz cujas colunas são os vectores de $S \subset V$, diz-se que S é **linearmente independente** se e só se $\mathcal{N}(A) = \{\mathbf{0}\}$.

Teorema 22. Seja A' uma matriz em escada de linhas.

(i) As colunas de A' que contêm pivots são linearmente independentes.

(ii) As linhas não nulas de A' são linearmente independentes.

(iii) O n^o de linhas independentes e o n^o de colunas independentes (de A') são ambos iguais à característica de A'.

Observação 15. **(i)** Assim, atendendo ao teorema anterior, a independência linear de $S = \{v_1, v_2, ..., v_k\} \subset V$ (espaço linear) pode ser decidida aplicando o método de eliminação à matriz A cujas colunas são os vectores de S, de modo a colocá-la em escada de linhas. Sendo A' essa matriz em escada, tem-se

$$\mathcal{N}(A) = \mathcal{N}(A') \quad (*).$$

Uma vez que as colunas de A' que contêm pivots são linearmente independentes então, devido a (*), as colunas de A nas posições correspondentes também serão linearmente independentes.

(ii) Em $\mathbb{R}$, quaisquer dois vectores são linearmente dependentes.

(iii) Em $\mathbb{R}^2$, dois vectores são linearmente independentes se não forem colineares.

(iv) Em $\mathbb{R}^3$, três vectores são linearmente independentes se não forem coplanares.

(v) Qualquer conjunto que contenha o vector nulo (elemento neutro) é linearmente dependente. Em particular, o conjunto $\{\mathbf{0}\}$, formado apenas pelo vector nulo, é linearmente dependente.

(vi) O conjunto vazio $\varnothing$ é linearmente independente.

Teorema 23. Sejam S_1 e S_2 dois subconjuntos finitos de um espaço linear, tais que $S_1 \subset S_2$.

(i) Se S_1 é linearmente dependente então S_2 também é linearmente dependente.

(ii) Se S_2 é linearmente independente então S_1 também é linearmente independente.

Observação 16. Sejam S_1 e S_2 dois subconjuntos finitos de um espaço linear, tais que $S_1 \subset S_2$.

(i) Se S_2 fôr linearmente dependente então S_1 tanto pode ser linearmente dependente como linearmente independente.

(ii) Se S_1 fôr linearmente independente então S_2 tanto pode ser linearmente dependente como linearmente independente.

Exemplo 20. Seja $S = \{(1,0,2),(2,0,4),(0,1,2)\}$. Tem-se

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 2 & 4 & 2 \end{bmatrix} \underset{-2L_1+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \underset{-2L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = A'.$$

Logo, como apenas existem dois pivots e portanto uma variável livre, as três colunas de A são linearmente dependentes, isto é, o conjunto S é linearmente dependente. O subconjunto de S:

$$\{(1,0,2),(2,0,4)\}$$

também é linearmente dependente. No entanto, uma vez que a 1^a e 3^a colunas de A são independentes pois correspondem às colunas da matriz em escada A' que contêm os pivots, o subconjunto de S:

$$\{(1,0,2),(0,1,2)\}$$

é linearmente independente.

Bases e dimensão de um espaço linear

Definição 33. Chama-se **base** de um espaço linear V a qualquer subconjunto $\mathcal{B}$ de V que verifique as duas condições:

(i) $\mathcal{B}$ gera V, isto é,

$$L(\mathcal{B}) = V.$$

(ii) $\mathcal{B}$ é linearmente independente.

Definição 34. Seja $\mathcal{B} = \{v_1, ..., v_k\}$ uma base ordenada de um espaço linear V e seja u um vector de V. Chamam-se **coordenadas** do vector u na base ordenada $\mathcal{B}$ aos escalares $\lambda_1, ..., \lambda_k$ da combinação linear:

$$u = \lambda_1 v_1 + ... + \lambda_k v_k.$$

Teorema 24. Seja V um espaço linear.

(i) Um conjunto $\mathcal{B}$ de vectores não nulos de V é uma base de V se e só se todo o vector de V puder ser escrito de modo único como combinação linear dos vectores de $\mathcal{B}$.

(ii) Se $\dim V = n$, então dados $u, w \in V$ e $\mathcal{B} = \{v_1, ..., v_n\}$ uma base ordenada de V, tem-se $u = w$ se e só se as coordenadas de u e de w na base $\mathcal{B}$ forem iguais.

Teorema 25. Qualquer espaço linear $V \neq \{\mathbf{0}\}$ tem pelo menos uma base.

Teorema 26. **(i)** Qualquer espaço linear $V \neq \{\mathbf{0}\}$ tem um n^o infinito de bases.

(ii) Seja $V \neq \{\mathbf{0}\}$ um espaço linear. Sejam $p, q \in \mathbb{N}$ tais que $\{u_1, ..., u_p\}$ é um conjunto gerador de V e $\{v_1, ..., v_q\}$ é um subconjunto de V linearmente independente. Então

$$p \geq q.$$

(iii) Todas as bases de um espaço linear $V \neq \{\mathbf{0}\}$ têm o mesmo n^o de vectores.

Dem. **(i)** Se $\mathcal{B} = \{u_1, ..., u_k\}$ fôr uma base de V então para cada $\alpha \neq 0$ o conjunto $\{\alpha u_1, ..., \alpha u_k\}$ é também uma base de V.

(ii) Suponhamos que $p < q$. Neste caso, como todos os vectores do conjunto $\{v_1, ..., v_q\}$ são não nulos por serem LI, poderíamos substituir sucessivamente os p vectores do conjunto $\{u_1, ..., u_p\}$ gerador de V por p vectores do conjunto $\{v_1, ..., v_q\}$, permitindo assim escrever cada vector do conjunto $\{v_{p+1}, ..., v_q\}$ como combinação linear do novo conjunto gerador de V: $\{v_1, ..., v_p\}$ e contrariando o facto dos vectores do conjunto $\{v_1, ..., v_q\}$ serem linearmente independentes.

Demonstração alternativa de (ii). Suponhamos que $p < q$. Como $\{u_1, ..., u_p\}$ gera V, para cada $j = 1, ..., q$ existem escalares $a_{1j}, ...a_{pj}$ tais que

$$v_j = \sum_{i=1}^{p} a_{ij} u_i.$$

Seja $A = (a_{ij})_{p\times q}$. Como $p < q$, o sistema homogéneo $A\alpha = \mathbf{0}$ é possível e indeterminado. Seja $\alpha = [\alpha_1 ... \alpha_q]^T \neq \mathbf{0}$ uma solução não nula de $A\alpha = \mathbf{0}$, isto é,

$$\mathbf{0} = \sum_{j=1}^{q} a_{ij}\alpha_j = \alpha_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{p1} \end{bmatrix} + ... + \alpha_q \begin{bmatrix} a_{1q} \\ \vdots \\ a_{pq} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{q} a_{1j}\alpha_j \\ \vdots \\ \sum_{j=1}^{q} a_{pj}\alpha_j \end{bmatrix}$$

com os α_j escalares não todos nulos. Por outro lado,

$$\begin{aligned} \sum_{j=1}^{q} \alpha_j v_j &= \sum_{j=1}^{q} \alpha_j \sum_{i=1}^{p} a_{ij} u_i = \sum_{i=1}^{p} \left(\sum_{j=1}^{q} a_{ij}\alpha_j \right) u_i = \\ &= \left(\sum_{j=1}^{q} a_{1j}\alpha_j \right) u_1 + ... + \left(\sum_{j=1}^{q} a_{pj}\alpha_j \right) u_p = \\ &= 0u_1 + ... + 0u_p = \mathbf{0} \end{aligned}$$

com os α_j não todos nulos, contrariando o facto dos vectores do conjunto $\{v_1, ..., v_q\}$ serem linearmente independentes.

(iii) Sendo $\{v_1, ..., v_q\}$ e $\{u_1, ..., u_p\}$ duas bases de V, por (i) tem-se $p \leq q$ e $q \leq p$. Logo $p = q$.

Definição 35. Chama-se **dimensão** de um espaço linear $V \neq \{\mathbf{0}\}$ ao n^o de vectores de uma base qualquer de V, e escreve-se $\dim V$. Se $V = \{\mathbf{0}\}$ então $\dim V = 0$ uma vez que o conjunto vazio $\varnothing$ é base de $\{\mathbf{0}\}$. Um espaço linear terá dimensão finita se uma sua base tiver um n^o finito de vectores.

Observação 17. A dimensão de um espaço linear, isto é, o n^o de elementos de uma sua base é igual ao n^o mínimo de vectores possam constituir um conjunto gerador desse espaço e é também igual ao n^o máximo de vectores que possam constituir um conjunto linearmente independente nesse espaço.

Exemplo 21. (i) O conjunto $\{1\}$ é uma base de $\mathbb{R}$, chamada base canónica ou natural de $\mathbb{R}$. Logo,

$$\dim \mathbb{R} = 1.$$

(ii) O conjunto $\{(1, 0), (0, 1)\}$ é uma base de $\mathbb{R}^2$, chamada base canónica ou natural de $\mathbb{R}^2$. Logo,

$$\dim \mathbb{R}^2 = 2.$$

(iii) O conjunto $\{(1,0,0),(0,1,0),(0,0,1)\}$ é uma base de $\mathbb{R}^3$, chamada base canónica ou natural de $\mathbb{R}^3$. Logo,

$$\dim \mathbb{R}^3 = 3.$$

(iv) Considerando $\mathbb{C}$ como corpo de escalares:

(a) o espaço linear $\mathbb{C}$ tem dimensão 1 sendo $\{1\}$ a base canónica de $\mathbb{C}$ uma vez que

$$a + bi = (a + bi)\, 1$$

(b) o espaço linear $\mathbb{C}^2$ tem dimensão 2 sendo $\{(1,0),(0,1)\}$ a base canónica de $\mathbb{C}^2$ uma vez que

$$(a + bi, c + di) = (a + bi)(1,0) + (c + di)(0,1).$$

(v) Considerando $\mathbb{R}$ como corpo de escalares:

(a) o espaço linear $\mathbb{C}$ tem dimensão 2 sendo $\{1, i\}$ a base canónica de $\mathbb{C}$ uma vez que

$$a + bi = a1 + bi$$

com $a, b \in \mathbb{R}$.

(b) o espaço linear $\mathbb{C}^2$ tem dimensão 4 sendo $\{(1,0),(i,0),(0,1),(0,i)\}$ a base canónica de $\mathbb{C}^2$ uma vez que

$$(a + bi, c + di) = a(1,0) + b(i,0) + c(0,1) + d(0,i)$$

com $a, b, c, d \in \mathbb{R}$.

(vi) O conjunto

$$\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$$

é uma base de $\mathcal{M}_{2\times 3}(\mathbb{R})$, chamada base canónica ou natural de $\mathcal{M}_{2\times 3}(\mathbb{R})$. Logo,

$$\dim \mathcal{M}_{2\times 3}(\mathbb{R}) = 6.$$

(vii) Tem-se

$$\dim \mathbb{R}^n = n \quad \text{e} \quad \dim \mathcal{M}_{m\times n}(\mathbb{R}) = mn.$$

(viii) O conjunto $\{1, t, t^2, ..., t^n\}$ é uma base de $\mathcal{P}_n$ (espaço linear de todos os polinómios reais de variável real e de grau menor ou igual a n), chamada base canónica ou natural de $\mathcal{P}_n$. Logo,

$$\dim \mathcal{P}_n = n + 1.$$

(ix) O conjunto $\{1, t, t^2, ...\}$ é uma base de $\mathcal{P}$ (espaço linear de todos os polinómios reais de variável real), chamada base canónica ou natural de $\mathcal{P}$. Logo,

$$\dim \mathcal{P} = \infty.$$

Definição 36. Seja A uma matriz do tipo $m \times n$. Chama-se característica de A e escreve-se $\operatorname{car} A$ à dimensão de $\mathcal{L}(A)$:

$$\operatorname{car} A = \dim \mathcal{L}(A)$$

Definição 37. Seja A uma matriz do tipo $m \times n$. Chama-se nulidade de A e escreve-se $\operatorname{nul} A$ à dimensão de $\mathcal{N}(A)$:

$$\operatorname{nul} A = \dim \mathcal{N}(A)$$

Teorema 27. Seja A uma matriz do tipo $m \times n$. Tem-se

$$\dim \mathcal{C}(A) = \dim \mathcal{L}(A) = \operatorname{car} A.$$

Dem. Suponhamos que $\operatorname{car} A = k$. Sendo A' a matriz $m \times n$ em escada (reduzida) de linhas, então A' tem exactamente k linhas não nulas. Sejam $R_1, ..., R_k$ essas linhas. Como

$$\mathcal{L}(A) = \mathcal{L}(A'),$$

então as linhas $L_1, ..., L_m$ de A podem ser expressas como combinações lineares das linhas $R_1, ..., R_k$, ou seja, existem escalares c_{ij}, com $i = 1, ..., m$ e $j = 1, ..., k$ tais que

$$\begin{array}{c} L_1 = c_{11}R_1 + ... + c_{1k}R_k \\ ... \\ L_m = c_{m1}R_1 + ... + c_{mk}R_k \end{array}$$

Para $i = 1, ..., m$, sejam a_{ij} e r_{ij} as componentes j das linhas L_i e R_i respectivamente. Assim, tem-se

$$\begin{array}{c} a_{1j} = c_{11}r_{1j} + ... + c_{1k}r_{kj} \\ ... \\ a_{mj} = c_{m1}r_{1j} + ... + c_{mk}r_{kj} \end{array}$$

ou seja, matricialmente,

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} = r_{1j} \begin{bmatrix} c_{11} \\ \vdots \\ c_{m1} \end{bmatrix} + ... + r_{kj} \begin{bmatrix} c_{1k} \\ \vdots \\ c_{mk} \end{bmatrix}.$$

Como $\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$ é a coluna j de A, a última igualdade mostra que os vectores

$$\begin{bmatrix} c_{11} \\ \vdots \\ c_{m1} \end{bmatrix}, ..., \begin{bmatrix} c_{1k} \\ \vdots \\ c_{mk} \end{bmatrix}$$

geram $\mathcal{C}(A)$. Logo, tem-se

$$\dim \mathcal{C}(A) \leq k = \dim \mathcal{L}(A).$$

Deste modo, substituindo A por A^T tem-se também

$$\underbrace{\dim \mathcal{C}\left(A^T\right)}_{=\dim \mathcal{L}(A)} \leq \underbrace{\dim \mathcal{L}\left(A^T\right)}_{=\dim \mathcal{C}(A)}.$$

Ou seja, tem-se

$$\dim \mathcal{C}\left(A\right) \leq \dim \mathcal{L}\left(A\right)$$

e

$$\dim \mathcal{L}\left(A\right) \leq \dim \mathcal{C}\left(A\right).$$

Isto é,

$$\dim \mathcal{C}\left(A\right) = \dim \mathcal{L}\left(A\right).$$

Observação 18. Atendendo ao teorema anterior tem-se

$$\text{car } A = \text{car } A^T$$

uma vez que

$$\text{car } A = \dim \mathcal{L}\left(A\right) = \dim \mathcal{C}\left(A\right) = \dim \mathcal{L}\left(A^T\right) = \text{car } A^T.$$

Teorema 28. Sejam V um espaço linear de dimensão finita e W um subespaço de V.

(i) Seja $S = \{u_1, ..., u_k\} \subset V$. Se S é linearmente independente então S será um subconjunto de uma base de V e ter-se-á $\dim V \geq k$.

(ii) Se $\dim V = n$, então quaisquer m vectores de V, com $m > n$, são linearmente dependentes.

(iii) Se $\dim V = n$, então nenhum conjunto com m vectores de V, em que $m < n$, pode gerar V.

(iv) O subespaço W tem dimensão finita e $\dim W \leq \dim V$.

(v) Se $\dim W = \dim V$, então $W = V$.

(vi) Se $\dim V = n$, então quaisquer n vectores de V linearmente independentes constituem uma base de V.

(vii) Se $\dim V = n$, então quaisquer n vectores geradores de V constituem uma base de V.

Exemplo 22. **(i)** Os seguintes conjuntos são **todos** os subespaços de $\mathbb{R}$: $\{0\}$ e $\mathbb{R}$.

(ii) Os seguintes conjuntos são **todos** os subespaços de $\mathbb{R}^2$:

$$\{(0,0)\}, \text{ todas as rectas que contêm a origem e } \mathbb{R}^2.$$

(iii) Os seguintes conjuntos são **todos** os subespaços de $\mathbb{R}^3$:

$$\{(0,0,0)\}\,, \text{ todas as rectas que contêm a origem,}$$

$$\text{todos os planos que contêm a origem e } \mathbb{R}^3.$$

Observação 19. O método de eliminação de Gauss permite determinar a dimensão e uma base quer para o espaço das linhas $\mathcal{L}(A)$ quer para o espaço das colunas $\mathcal{C}(A)$ de uma matriz A. Seja A' a matriz em escada que se obtem de A por aplicação do método de eliminação de Gauss. Então,

(i) Uma base para $\mathcal{L}(A)$ será formada pelas linhas não nulas de A'.

(ii) Uma base para $\mathcal{C}(A)$ será formada pelas colunas de A que correspondem às posições das colunas de A' que contêm os pivots.

Exemplo 23. Seja $A = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 2 & 3 & 3 \\ -6 & -3 & 1 & 1 \end{bmatrix}$. Tem-se

$$A = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 2 & 3 & 3 \\ -6 & -3 & 1 & 1 \end{bmatrix} \underset{\substack{-2L_1+L_2 \to L_2 \\ 3L_1+L_3 \to L_3}}{\longrightarrow} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 4 & 4 \end{bmatrix} \underset{-4L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = A'.$$

Logo, $\{(2,1,1,1),(0,0,1,1)\}$ é uma base de $\mathcal{L}(A)$ e $\{(2,4,-6),(1,3,1)\}$ é uma base de $\mathcal{C}(A)$. Assim,

$$\dim \mathcal{L}(A) = 2 = \dim \mathcal{C}(A)$$

e

$$\mathcal{L}(A) = L\left(\{(2,1,1,1),(0,0,1,1)\}\right), \ \ \mathcal{C}(A) = L\left(\{(2,4,-6),(1,3,1)\}\right).$$

Por outro lado,

$$\mathcal{N}(A') = \left\{ (x,y,z,w) \in \mathbb{R}^4 : A' \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\} =$$

$$= \{(x,-2x,-w,w) : x,w \in \mathbb{R}\} = L\left(\{(1,-2,0,0),(0,0,-1,1)\}\right).$$

Como o conjunto $\{(1,-2,0,0),(0,0,-1,1)\}$ é linearmente independente e gera $\mathcal{N}(A')$ então é uma base de $\mathcal{N}(A')$. Finalmente, uma vez que $\mathcal{N}(A) = \mathcal{N}(A')$, o conjunto

$$\{(1,-2,0,0),(0,0,-1,1)\}$$

é uma base de $\mathcal{N}(A)$ e portanto $\dim \mathcal{N}(A) = 2$, com

$$\mathcal{N}(A) = L\left(\{(1,-2,0,0),(0,0,-1,1)\}\right).$$

Exemplo 24. Seja

$$S = \{1, 2, -1), (2, 1, 1), (-1, -2, 1), (0, 1, 0)\} \subset \mathbb{R}^3.$$

Determinemos uma base para $L(S)$.

Considerando a matriz cujas colunas são os vectores de S, tem-se

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & -2 & 1 \\ -1 & 1 & 1 & 0 \end{bmatrix} \underset{\substack{-2L_1+L_2 \to L_2 \\ L_1+L_3 \to L_3}}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & -3 & 0 & 1 \\ 0 & 3 & 0 & 0 \end{bmatrix} \underset{L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & -3 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Logo, $S' = \{1, 2, -1), (2, 1, 1), (0, 1, 0)\}$ é uma base de $L(S)$. Como $\dim \mathbb{R}^3 = 3$, então tem-se mesmo: $L(S) = \mathbb{R}^3$ e S' é uma base de $\mathbb{R}^3$.

Resolução alternativa: Considerando a matriz cujas linhas são os vectores de S, tem-se

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ -1 & -2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underset{\substack{-2L_1+L_2 \to L_2 \\ L_1+L_3 \to L_3}}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \underset{L_3 \leftrightarrow L_4}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \underset{\frac{1}{3}L_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -3 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Logo, $S' = \{1, 2, -1), (0, -3, 3), (0, 0, 1)\}$ é uma base de $L(S)$. Como $\dim \mathbb{R}^3 = 3$, então tem-se mesmo: $L(S) = \mathbb{R}^3$ e S' é uma base de $\mathbb{R}^3$.

Exemplo 25. Seja

$$S_{a,b} = \{1, 0, 1), (0, 1, a), (1, 1, b), (1, 1, 1)\} \subset \mathbb{R}^3.$$

Determinemos os valores dos parâmetros a e b para os quais $S_{a,b}$ não gere $\mathbb{R}^3$.

Considerando a matriz cujas colunas são os vectores de S, tem-se

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & a & b & 1 \end{bmatrix} \underset{-L_1+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & a & b-1 & 0 \end{bmatrix} \underset{-aL_2+L_3 \to L_3}{\longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & b-a-1 & -a \end{bmatrix}.$$

Logo, $S_{a,b}$ não gera $\mathbb{R}^3$ se e só se $b - a - 1 = 0$ e $-a = 0$, isto é, se e só se $a = 0$ e $b = 1$.

Teorema 29. (i) Seja $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. As colunas de A geram $\mathbb{R}^m$ se e só se

$$\operatorname{car} A = m.$$

(ii) Seja $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. As colunas de A são linearmente independentes se e só se

$$\operatorname{car} A = n.$$

(iii) Seja $A \in \mathcal{M}_{n \times n}(\mathbb{R})$. A matriz A é invertível se e só se as colunas de A (ou as linhas de A) formarem uma base de $\mathbb{R}^n$. No caso de A ser invertível tem-se

$$\mathcal{C}(A) = \mathcal{L}(A) = \mathbb{R}^n.$$

Teorema 30. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e considere o sistema de equações lineares $Au = b$.

(i) O sistema $Au = b$ é **impossível** (não tem solução) se e só se $b \notin \mathcal{C}(A)$, isto é, se e só se $\operatorname{car} A < \operatorname{car}[A \mid b]$.

(ii) O sistema $Au = b$ é **possível e indeterminado** (tem um n^o infinito de soluções) se e só se $b \in \mathcal{C}(A)$ e as colunas de A forem linearmente dependentes, isto é, se e só se

$$\operatorname{car} A = \operatorname{car}[A \mid b] < n,$$

isto é, se e só se

$$\operatorname{car} A = \operatorname{car}[A \mid b] \quad \text{e} \quad \operatorname{nul} A \neq 0.$$

(iii) O sistema $Au = b$ é **possível e determinado** (tem uma única solução) se e só se $b \in \mathcal{C}(A)$ e as colunas de A forem linearmente independentes, isto é, se e só se

$$\operatorname{car} A = \operatorname{car}[A \mid b] = n,$$

isto é, se e só se

$$\operatorname{car} A = \operatorname{car}[A \mid b] \quad \text{e} \quad \operatorname{nul} A = 0.$$

Teorema 31. Sejam W_1 e W_2 dois subespaços de dimensão finita de um espaço linear V. Então,

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2),$$

Dem. Sejam

$$n = \dim W_1, \qquad m = \dim W_2 \qquad \text{e} \qquad k = \dim(W_1 \cap W_2).$$

Se $k = 0$ a igualdade do teorema é imediata. Se $k \neq 0$, seja $\{w_1, ..., w_k\}$ uma base de $W_1 \cap W_2$. Sejam $u_{k+1}, ..., u_n \in W_1$ tais que

$$\{w_1, ..., w_k, u_{k+1}, ..., u_n\}$$

é uma base de W_1. Sejam $v_{k+1}, ..., v_m \in W_2$ tais que

$$\{w_1, ..., w_k, v_{k+1}, ..., v_m\}$$

é uma base de W_2. Vejamos que

$$\mathcal{B} = \{w_1, ..., w_k, u_{k+1}, ..., u_n, v_{k+1}, ..., v_m\}$$

é uma base de $W_1 + W_2$.

Seja $w \in W_1 + W_2$. Existem $u \in W_1$ e $v \in W_2$ tais que $w = u + v$. Ou seja, existem escalares (únicos) $\alpha_1, ..., \alpha_n$ e $\beta_1, ..., \beta_m$ tais que

$$w = u + v = \sum_{i=1}^{k} (\alpha_i + \beta_i) w_i + \sum_{j=k+1}^{n} \alpha_j u_j + \sum_{l=k+1}^{m} \beta_l v_l$$

pelo que $\mathcal{B}$ gera $W_1 + W_2$.

Sejam $\gamma_1, ..., \gamma_n, \xi_{k+1}, ..., \xi_m$ $n + m - k$ escalares tais que

$$\mathbf{0} = \sum_{i=1}^{k} \gamma_i w_i + \sum_{j=k+1}^{n} \gamma_j u_j + \sum_{l=k+1}^{m} \xi_l v_l,$$

isto é,

$$\sum_{l=k+1}^{m} \xi_l v_l = -\left(\sum_{i=1}^{k} \gamma_i w_i + \sum_{j=k+1}^{n} \gamma_j u_j\right) \in W_1,$$

ou seja

$$\sum_{l=k+1}^{m} \xi_l v_l \in W_1 \cap W_2.$$

Atendendo a que $\{w_1, ..., w_k\}$ é base de $W_1 \cap W_2$, existem escalares $\eta_1, ..., \eta_k$ tais que

$$\sum_{l=k+1}^{m} \xi_l v_l = \sum_{i=1}^{k} \eta_i w_i,$$

isto é,

$$\sum_{i=1}^{k} \eta_i w_i + \sum_{l=k+1}^{m} (-\xi_l)\, v_l = \mathbf{0}.$$

Como $\{w_1, ..., w_k, v_{k+1}, ..., v_m\}$ é uma base de W_2, tem-se

$$\eta_1 = ... = \eta_k = \xi_{k+1} = ... = \xi_m = 0.$$

Logo

$$\mathbf{0} = \sum_{i=1}^{k} \gamma_i w_i + \sum_{j=k+1}^{n} \gamma_j u_j + \sum_{l=k+1}^{m} \xi_l v_l = \sum_{i=1}^{k} \gamma_i w_i + \sum_{j=k+1}^{n} \gamma_j u_j.$$

Assim, como $\{w_1, ..., w_k, u_{k+1}, ..., u_n\}$ é uma base de W_1, tem-se $\gamma_1 = ... = \gamma_n = 0$. Deste modo, como

$$\gamma_1 = ... = \gamma_n = \xi_{k+1} = ... = \xi_m = 0$$

então o conjunto $\mathcal{B}$ é linearmente independente.

Teorema 32. Seja A uma matriz do tipo $m \times n$. Tem-se

$$\mathbb{R}^n = \mathcal{N}(A) \oplus \mathcal{L}(A).$$

Dem. Como

$$\dim\left(\mathcal{N}(A) + \mathcal{L}(A)\right) = \dim \mathcal{N}(A) + \dim \mathcal{L}(A) - \dim\left(\mathcal{N}(A) \cap \mathcal{L}(A)\right) =$$

$$= \operatorname{nul} A + \operatorname{car} A - 0 = n$$

e $\mathcal{N}(A) + \mathcal{L}(A)$ é um subespaço de $\mathbb{R}^n$ então

$$\dim \mathbb{R}^n = \dim \mathcal{N}(A) + \dim \mathcal{L}(A)$$

e

$$\mathbb{R}^n = \mathcal{N}(A) \oplus \mathcal{L}(A).$$

Determinantes

Definição 38. Dados os números naturais $1, 2, ..., n$ chama-se **permutação** desses n números a qualquer lista em que os mesmos sejam apresentados por ordem arbitrária.

Definição 39. Seja $(i_1 i_2 ... i_n)$ uma permutação dos números naturais $1, 2, ..., n$. Diz-se que um par $(i_j i_k)$ é uma **inversão** quando $(j-k)(i_j - i_k) < 0$ (isto é, quando i_j e i_k aparecerem na permutação por ordem decrescente).

Definição 40. Uma permutação $(i_1 i_2 ... i_n)$ diz-se **par** (**ímpar**) quando o n^o máximo de inversões incluídas fôr par (ímpar).

Exemplo 26. A permutação (21453) é ímpar pois o n^o máximo de inversões nela incluídas é ímpar: $(21), (43)$ e (53).

Definição 41. Seja A uma matriz do tipo $n \times n$. Chama-se **determinante** de A, e escreve-se $|A|$ ou $\det A$, o número que se obtém do seguinte modo:

(i) Formam-se todos os produtos possíveis de n factores em que não intervenha mais do que um elemento da mesma linha e da mesma coluna de A.

(ii) Afecta-se cada produto do sinal $+$ ou do sinal $-$ conforme as permutações (dos números naturais $1, 2, ..., n$) que figuram nos índices de linha e de coluna tenham a mesma paridade ou não.

(iii) Somam-se as parcelas obtidas.

Em resumo: fixando, por exemplo, a permutação $(i_1 i_2 ... i_n)$ de $1, 2, ..., n$

$$|A| = \sum_{\substack{(j_1 j_2 ... j_n) \\ \text{permutação de } 1,2,...,n}} (-1)^{\sigma} a_{i_1 j_1} a_{i_2 j_2} ... a_{i_n j_n},$$

em que

$$\sigma = \begin{cases} 0 & \text{se } (i_1 i_2 ... i_n) \text{ e } (j_1 j_2 ... j_n) \text{ têm a mesma paridade} \\ 1 & \text{se } (i_1 i_2 ... i_n) \text{ e } (j_1 j_2 ... j_n) \text{ têm paridade diferente.} \end{cases}$$

Observação 20. Pode ainda escrever-se

$$|A| = \sum_{\substack{(j_1 j_2 ... j_n) \\ \text{permutação de } 1,2,...,n}} (-1)^\sigma a_{1j_1} a_{2j_2} ... a_{nj_n} \quad \text{onde} \quad \sigma = \begin{cases} 0 & \text{se } (j_1 j_2 ... j_n) \text{ é par} \\ 1 & \text{se } (j_1 j_2 ... j_n) \text{ é ímpar.} \end{cases}$$

ou

$$|A| = \sum_{\substack{(i_1 i_2 ... i_n) \\ \text{permutação de } 1,2,...,n}} (-1)^\sigma a_{i_1 1} a_{i_2 2} ... a_{i_n n} \quad \text{onde} \quad \sigma = \begin{cases} 0 & \text{se } (i_1 i_2 ... i_n) \text{ é par} \\ 1 & \text{se } (i_1 i_2 ... i_n) \text{ é ímpar.} \end{cases}$$

Teorema 33. (i) Se A é do tipo 2×2, então

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

(ii) Se A é do tipo 3×3, então

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Observação 21. Se A é uma matriz do tipo $n \times n$ então $|A|$ tem $n!$ parcelas.

Exemplo 27. (i)

$$\begin{vmatrix} 1 & -1 \\ 2 & -2 \end{vmatrix} = 1(-2) - (-1)2 = 0.$$

(ii)

$$\begin{vmatrix} 1 & 2 & 1 \\ 3 & -1 & 2 \\ 2 & 1 & -3 \end{vmatrix} = 1(-1)(-3) + 3 + 8 - 1(-1)2 - 6(-3) - 2 = 32.$$

Definição 42. Seja $A = (a_{ij})$ uma matriz do tipo $n \times n$, com $n > 1$. Seja A_{ij} a matriz do tipo $(n-1) \times (n-1)$ que se obtem de A suprimindo a linha i e a coluna j de A. Chama-se a A_{ij} o **menor-**ij da matriz A.

Teorema 34. (Fórmula de Laplace.) Seja A uma matriz do tipo $n \times n$, com $n > 1$. Tem-se

$$\det A = \sum_{j=1}^{n} a_{ij}(-1)^{i+j} \det A_{ij}, \qquad \text{com } i \in \{1, ..., n\} \text{ fixo.}$$

Observação 22. Seja A uma matriz do tipo $n \times n$, com $n > 1$. Tem-se

$$\det A = \sum_{i=1}^{n} a_{ij}(-1)^{i+j} \det A_{ij}, \qquad \text{com } j \in \{1, ..., n\} \text{ fixo.}$$

Exemplo 28.

$$\begin{vmatrix} 1 & 0 & -2 & 3 \\ 2 & 1 & -1 & 4 \\ 0 & -1 & 0 & -2 \\ 1 & 0 & -2 & -3 \end{vmatrix} = (-1)(-1)^{3+2} \begin{vmatrix} 1 & -2 & 3 \\ 2 & -1 & 4 \\ 1 & -2 & -3 \end{vmatrix} + (-2)(-1)^{3+4} \begin{vmatrix} 1 & 0 & -2 \\ 2 & 1 & -1 \\ 1 & 0 & -2 \end{vmatrix} =$$

$$= (-1)(-3) + (-2)4 + 2(-2)3 - (-1)3 - (-2)2(-3) - 4(-2) + 2\left[(-2) - (-2)\right] = -18.$$

Teorema 35. Sejam A e B matrizes do tipo $n \times n$. Seja λ um escalar.

(i) $\det\left(A^T\right) = \det A$.

(ii) Se A fôr uma matriz diagonal, triangular superior ou triangular inferior então o determinante de A é igual ao produto dos elementos da diagonal principal de A.

(iii) Se A tiver uma linha (ou coluna) nula então $\det A = 0$.

(iv) Se B fôr obtida de A trocando duas linhas (ou colunas) de A então $\det B = -\det A$.

(v) Sendo B, A_1 e A_2 matrizes do tipo $n \times n$ com as $n-1$ linhas (colunas): $1, 2, ..., i-1, i+1, ..., n$ iguais, se a linha (coluna) i de B fôr obtida somando as linhas (colunas) i de A_1 e de A_2 então $\det B = \det A_1 + \det A_2$.

(vi) Sendo B fôr obtida de A multiplicando uma linha (ou coluna) de A por um escalar λ então $\det B = \lambda \det A$.

(vii) Se duas linhas (ou colunas) de A forem iguais então $\det A = 0$.

(viii) Se B fôr obtida de A somando a uma linha (ou coluna) de A um múltiplo escalar λ de uma outra linha (ou coluna) de A então $\det B = \det A$.

(ix) $\det(\lambda A) = \lambda^n \det A$.

(x) $\det A \neq 0 \Leftrightarrow A$ é invertível.

(xi) $\det(AB) = \det A \det B$.

(xii) $\det(A_1 A_2 ... A_l) = \det A_1 \det A_2 ... \det A_l$, onde $A_1, A_2, ..., A_l$ são l $(l \in \mathbb{N})$ matrizes do tipo $n \times n$.

(xiii) Se A fôr invertível, $\det(A^{-1}) = \dfrac{1}{\det A}$.

(xiv) $\det(AB) = 0 \Leftrightarrow (\det A = 0 \ \text{ ou } \ \det B = 0)$.

(xv) $\det(AB) = \det(BA)$.

Exemplo 29.

$$\begin{vmatrix} 9 & 7 & 5 & 3 & 1 \\ 7 & 7 & 5 & 3 & 1 \\ 5 & 5 & 5 & 3 & 1 \\ 3 & 3 & 3 & 3 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 3 & 3 & 3 & 1 \\ 5 & 5 & 5 & 3 & 1 \\ 7 & 7 & 5 & 3 & 1 \\ 9 & 7 & 5 & 3 & 1 \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 3 & 3 & 3 \\ 1 & 3 & 5 & 5 & 5 \\ 1 & 3 & 5 & 7 & 7 \\ 1 & 3 & 5 & 7 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 & 2 \\ 0 & 2 & 4 & 4 & 4 \\ 0 & 2 & 4 & 6 & 6 \\ 0 & 2 & 4 & 6 & 8 \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 & 4 \\ 0 & 0 & 2 & 4 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 2 \end{vmatrix} = 2^4 = 16.$$

Observação 23. **(i)** Sendo A e B matrizes do tipo $n \times n$, em **geral**:

$$|A + B| \neq |A| + |B| \qquad \text{e} \qquad |A - B| \neq |A| - |B|.$$

Por exemplo, se n é par, $A = I$ e $B = -I$, tem-se

$$|A + B| = 0 \neq 2 \underset{n \text{ é par}}{=} 1 + (-1)^n = |A| + |B|.$$

(ii) Sendo A uma matriz do tipo $n \times n$, se fixarmos $n - 1$ linhas (colunas), o determinante de A é uma função linear em relação à linha (coluna) não fixada.

Definição 43. Seja $A = (a_{ij})$ uma matriz do tipo $n \times n$, com $n > 1$. Seja A_{ij} o menor-ij da matriz A. Chama-se a $(-1)^{i+j} \det A_{ij}$ o **cofactor-**ij da matriz A e à matriz $\operatorname{cof} A = ((-1)^{i+j} \det A_{ij})$ do tipo $n \times n$, com $n > 1$, a matriz dos cofactores de A.

Teorema 36. Para qualquer matriz A do tipo $n \times n$, com $n > 1$, tem-se

$$A\,(\operatorname{cof} A)^T = (\operatorname{cof} A)^T A = (\det A)\, I.$$

Se $\det A \neq 0$ então A é invertível e

$$A^{-1} = \frac{1}{\det A} (\operatorname{cof} A)^T = \left(\underbrace{\frac{1}{\det A} (-1)^{j+i} \det A_{ji}}_{\text{entrada } (i,j) \text{ de } A^{-1}} \right)_{n \times n}.$$

Exemplo 30. **(i)** Seja $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R})$ tal que $\det A \neq 0$. Então A é invertível e

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Note que $ad - bc = \det A$.

(ii) Podemos usar o teorema anterior para calcular não só a inversa de uma matriz (invertível) mas também (e sobretudo) entradas concretas dessa inversa. Seja

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

A entrada $(1, 2)$ da matriz A^{-1} é dada por

$$(A^{-1})_{12} = \frac{1}{\det A} \left((\operatorname{cof} A)^T \right)_{12} = \frac{1}{\det A} \left((-1)^{2+1} \det A_{21} \right) = \frac{1}{-12} \left(-\det \left(\begin{bmatrix} 0 & 3 \\ 8 & 9 \end{bmatrix} \right) \right) = -2.$$

Note que apesar da entrada $(1, 2)$ de A ser nula, a entrada $(1, 2)$ de A^{-1} não é nula.

(iii) Para calcular A^{-1} a partir do teorema anterior, é preciso calcular $(\operatorname{cof} A)^T$. Assim, usando por exemplo A da alínea anterior, tem-se

$$\operatorname{cof} A = \begin{bmatrix} \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} & -\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} & \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} \\ -\begin{vmatrix} 0 & 3 \\ 8 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 0 \\ 7 & 8 \end{vmatrix} \\ \begin{vmatrix} 0 & 3 \\ 5 & 6 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 4 & 5 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -3 & 6 & -3 \\ 24 & -12 & -8 \\ -15 & 6 & 5 \end{bmatrix}$$

pelo que

$$(\operatorname{cof} A)^T = \begin{bmatrix} -3 & 24 & -15 \\ 6 & -12 & 6 \\ -3 & -8 & 5 \end{bmatrix}$$

e assim

$$A^{-1} = \frac{1}{\det A} (\operatorname{cof} A)^T = \frac{1}{-12} \begin{bmatrix} -3 & 24 & -15 \\ 6 & -12 & 6 \\ -3 & -8 & 5 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & -2 & \frac{5}{4} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ \frac{1}{4} & \frac{2}{3} & -\frac{5}{12} \end{bmatrix}.$$

De facto

$$\begin{bmatrix} \frac{1}{4} & -2 & \frac{5}{4} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ \frac{1}{4} & \frac{2}{3} & -\frac{5}{12} \end{bmatrix} \begin{bmatrix} 1 & 0 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Teorema 37. (Regra de Cramer.) Seja A uma matriz do tipo $n \times n$ tal que A é invertível. Então a única solução do sistema de equações lineares $AX = B$ é dada por

$$X = A^{-1}B = \frac{1}{\det A} (\operatorname{cof} A)^T B.$$

Isto é, sendo $X = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T$ e $B = \begin{bmatrix} b_1 & \ldots & b_n \end{bmatrix}^T$ tem-se, para $i = 1, \ldots, n$,

$$x_i = \frac{1}{\det A} \sum_{k=1}^{n} (-1)^{k+i} \det A_{ki} b_k = \frac{\det C_i}{\det A},$$

onde C_i é a matriz obtida de A substituindo a coluna i de A pela matriz coluna B dos termos independentes.

Exemplo 31. O sistema de equações lineares

$$\begin{cases} y + 2z = 8 \\ 4x + 2y - z = 7 \\ x - z = 1 \end{cases}$$

pode ser resolvido usando a regra de Cramer:

$$x = \frac{\begin{vmatrix} 8 & 1 & 2 \\ 7 & 2 & -1 \\ 1 & 0 & -1 \end{vmatrix}}{\begin{vmatrix} 0 & 1 & 2 \\ 4 & 2 & -1 \\ 1 & 0 & -1 \end{vmatrix}} = 14, \quad y = \frac{\begin{vmatrix} 0 & 8 & 2 \\ 4 & 7 & -1 \\ 1 & 1 & -1 \end{vmatrix}}{\begin{vmatrix} 0 & 1 & 2 \\ 4 & 2 & -1 \\ 1 & 0 & -1 \end{vmatrix}} = -18 \quad \text{e} \quad z = \frac{\begin{vmatrix} 0 & 1 & 8 \\ 4 & 2 & 7 \\ 1 & 0 & 1 \end{vmatrix}}{\begin{vmatrix} 0 & 1 & 2 \\ 4 & 2 & -1 \\ 1 & 0 & -1 \end{vmatrix}} = 13.$$

Valores próprios e vectores próprios de uma matriz. Diagonalização.

Definição 44. Seja A uma matriz $n \times n$. Chama-se ao polinómio

$$p(\lambda) = \det(A - \lambda I)$$

o **polinómio característico** da matriz A. Este polinómio tem grau n, o coeficiente do termo de grau n é $(-1)^n$, o coeficiente do termo de grau $n-1$ é $(-1)^{n-1} \operatorname{tr} A$ e o termo constante é $p(0) = \det A$.

Definição 45. Seja A uma matriz $n \times n$. Chama-se **valor próprio** de A a qualquer escalar λ tal que $A - \lambda I$ seja não invertível, isto é, tal que $\det(A - \lambda I) = 0$. Ao conjunto de todos os valores próprios de A chama-se **espectro** de A. À multiplicidade de λ como raíz do polinómio $\det(A - \lambda I)$ chama-se **multiplicidade algébrica** de λ e denota-se por $m_a(\lambda)$. Chama-se **vector próprio** de A, associado ao valor próprio λ de A, a qualquer vector não nulo v que verifique

$$(A - \lambda I)v = \mathbf{0},$$

isto é, a qualquer vector

$$v \in \mathcal{N}(A - \lambda I) \setminus \{\mathbf{0}\}.$$

Teorema 38. Seja A uma matriz $n \times n$. O escalar 0 é valor próprio de A se e só se A fôr não invertível. Isto é, a matriz A é invertível se e só se 0 não fôr valor próprio de A.

Teorema 39. Seja A uma matriz $n \times n$. Então o polinómio característico de A pode ser escrito na forma:

$$p(\lambda) = \det(A - \lambda I) = (\lambda_1 - \lambda)^{m_1} (\lambda_2 - \lambda)^{m_2} \cdots (\lambda_k - \lambda)^{m_k},$$

onde $\lambda_1, \lambda_2, \ldots, \lambda_k$ são os valores próprios distintos de A e $m_1, m_2, \ldots, m_k$ são tais que $m_1 + m_2 + \cdots + m_k = n$.

Definição 46. Se

$$p(\lambda) = \det(A - \lambda I) = (\lambda_1 - \lambda)^{m_1} (\lambda_2 - \lambda)^{m_2} \cdots (\lambda_k - \lambda)^{m_k},$$

onde $\lambda_1, \lambda_2, \ldots, \lambda_k$ são os valores próprios distintos de A, aos expoentes $m_1, m_2, \ldots, m_k$ chamam-se as **multiplicidades algébricas** desses valores próprios respectivamente. Escreve-se

$$m_a(\lambda_k) = m_k.$$

Teorema 40. Seja A uma matriz $n \times n$, com os valores próprios $\lambda_1, \lambda_2, \ldots, \lambda_n$ (repetidos de acordo com a respectiva multiplicidade algébrica). Então, atendendo à alínea anterior e à definição anterior tem-se

$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n \qquad \text{e} \qquad \operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

Definição 47. Sejam A e B matrizes $n \times n$. As matrizes A e B dizem-se **semelhantes** se existir uma matriz S invertível tal que

$$B = SAS^{-1}.$$

Teorema 41. Sejam A e B matrizes $n \times n$. Se A e B forem semelhantes **então** A e B têm o(a) mesmo(a)**:**

(i) determinante; **(ii)** característica; **(iii)** nulidade; **(iv)** traço;

(v) polinómio característico, e portanto têm os mesmos valores próprios com as mesmas multiplicidades algébricas e geométricas.

Dem. (Matrizes semelhantes têm o mesmo polinómio característico.)

$$\begin{aligned}
\det(B - \lambda I) &= \det(SAS^{-1} - \lambda I) = \det(SAS^{-1} - \lambda SS^{-1}) = \\
&= \det(S(A - \lambda I)S^{-1}) = \det S \det(A - \lambda I) \det\left(S^{-1}\right) = \\
&= \det S \det(A - \lambda I) \frac{1}{\det S} = \det(A - \lambda I).
\end{aligned}$$

Teorema 42. **(i)** Seja A uma matriz $n \times n$. Se A tiver valores próprios $\lambda_1, ..., \lambda_k$ distintos dois a dois e se para cada $i = 1, ..., k$ considerarmos o conjunto S_i dos vectores próprios de A linearmente independentes e associados a λ_i, então $S_1 \cup ... \cup S_k$ é um conjunto linearmente independente.

(ii) Seja A uma matriz $n \times n$. Tem-se

$$m_g(\lambda_i) \leq m_a(\lambda_i),$$

para qualquer valor próprio λ_i de A.

Dem. **(i)** Vejamos que a afirmação é válida para $k = 2$. O caso geral prova-se por indução. Sejam λ_1 e λ_2 dois valores próprios distintos e sejam $S_1 = \{u_1, ..., u_r\}$ e $S_2 = \{v_1, ..., v_s\}$ dois conjuntos de vectores próprios de A linearmente independentes e associados respectivamente a λ_1 e a λ_2. Suponhamos que se tinha

$$\alpha_1 u_1 + \cdots + \alpha_r u_r + \beta_1 v_1 + \cdots + \beta_s v_s = \mathbf{0}. \text{ (*)}$$

Logo

$$\begin{aligned}
\mathbf{0} &= A(\alpha_1 u_1 + \cdots + \alpha_r u_r + \beta_1 v_1 + \cdots + \beta_s v_s) = \\
&= \alpha_1 \lambda_1 u_1 + \cdots + \alpha_r \lambda_1 u_r + \beta_1 \lambda_2 v_1 + \cdots + \beta_s \lambda_2 v_s. \text{ (**)}
\end{aligned}$$

Multiplicando (*) por λ_1 e subtraindo a (**) obtem-se

$$\beta_1(\lambda_2 - \lambda_1) v_1 + \cdots + \beta_s(\lambda_2 - \lambda_1) v_s = \mathbf{0},$$

e atendendo a que $\lambda_1 \neq \lambda_2$ e ao facto de S_2 ser linearmente independente, conclui-se que $\beta_1 = \cdots = \beta_s = 0$. Finalmente, como S_1 é linearmente independente, então $\alpha_1 = \cdots = \alpha_r = 0$ e deste modo $S_1 \cup S_2$ é um conjunto linearmente independente.

(ii) Seja λ_i um qualquer valor próprio de A. Seja $r = m_g(\lambda_i) = \dim \mathcal{N}(A - \lambda_i I)$. Seja $\{u_1, ..., u_r\}$ uma base de $\mathcal{N}(A - \lambda_i I)$. Seja $\{u_1, ..., u_r, u_{r+1}, ..., u_n\}$ uma base de $\mathbb{R}^n$ (ou de $\mathbb{C}^n$). Considere-se a matriz invertível $S^{-1} = [u_1 ... u_r u_{r+1} ... u_n]$. Tem-se

$$SAS^{-1} = \begin{bmatrix} \lambda_i I_{r \times r} & * \\ \mathbf{0}_{(n-r)\times r} & ** \end{bmatrix}.$$

Logo, como SAS^{-1} e A têm o mesmo polinómio característico, então λ_i é uma raíz do polinómio característico de A com multiplicidade algébrica pelo menos igual a r.

Definição 48. Seja A uma matriz $n \times n$. Se existir uma matriz P^{-1} invertível tal que

$$D \;=\; PAP^{-1},$$

com D **matriz diagonal**, então diz-se que A é uma **matriz diagonalizável** e que P^{-1} é a **matriz diagonalizante**. No caso de A ser uma matriz diagonal, a matriz diagonalizante é a matriz identidade.

Teorema 43. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. A matriz A é diagonalizável se e só se existir uma base $\mathcal{B}_{vp}$ de $\mathbb{R}^n$ apenas constituída por vectores próprios de A. Neste caso, as entradas da diagonal principal da matriz diagonal D serão os valores próprios de A apresentados pela ordem dos vectores próprios correspondentes na base ordenada $\mathcal{B}_{vp}$. Além disso, a matriz P^{-1} será a matriz cujas colunas serão os vectores próprios de A, da base $\mathcal{B}_{vp}$ de $\mathbb{R}^n$ dispostos pela mesma ordem, tendo-se

$$D = PAP^{-1}.$$

O mesmo se aplica a $\mathbb{C}^n$.

Teorema 44. Seja A uma matriz $n \times n$. Sendo $\lambda_1, ..., \lambda_k$ os valores próprios distintos de A, então as afirmações seguintes são equivalentes:

(i) A é diagonalizável.

(ii) A tem n vectores próprios linearmente independentes.

(iii) $\sum_{i=1}^{k} m_g(\lambda_i) = n$.

(iv) $m_g(\lambda_i) = m_a(\lambda_i)$ para todo o $i = 1, ..., k$..

Dem. **(i) $\Leftrightarrow$(iii)** Sejam $\lambda_1, ..., \lambda_k$ os valores próprios de A distintos dois a dois.

($\Rightarrow$) Suponhamos que A é diagonalizável. Então A terá n vectores próprios linearmente independentes. Suponhamos que l_i dos vectores próprios de A estão associados ao valor próprio λ_i. Logo, para cada $i = 1, ..., k$

$$\dim \mathcal{N}(A - \lambda_i I) \geq l_i.$$

Seja

$$r = \dim \mathcal{N}(A - \lambda_1 I) + ... + \dim \mathcal{N}(A - \lambda_k I).$$

Então

$$r \geq l_1 + ... + l_k = n.$$

Para cada $i = 1, ..., k$ seja S_i uma base de $\mathcal{N}(A - \lambda_i I)$. Logo $S_1 \cup ... \cup S_k$ é um conjunto de r vectores linearmente independentes, pelo que se tem $r \leq n$. Logo $r = n$.

($\Leftarrow$) Suponhamos que $n = \dim \mathcal{N}(A - \lambda_1 I) + ... + \dim \mathcal{N}(A - \lambda_k I)$. Para cada $i = 1, ..., k$ sendo $m_i = \dim \mathcal{N}(A - \lambda_i I)$, existirá então um conjunto S_i formado por m_i vectores próprios de A linearmente independentes associados ao valor próprio λ_i. Assim, conclui-se que $S_1 \cup ... \cup S_k$ é um conjunto de n vectores próprios de A linearmente independentes, sendo deste modo A diagonalizável.

Observação 24. (i) Se todos os valores próprios de A forem raízes simples do polinómio característico, então A é diagonalizável.

(ii) Se $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ então A é é diagonalizável se e só se:

$$\mathbb{R}^n = \mathcal{N}(A - \lambda_1 I) \oplus ... \oplus \mathcal{N}(A - \lambda_k I).$$

(iii) No caso de se ter $D = PAP^{-1}$, com P^{-1} invertível e D matriz diagonal, tem-se, para $k \in \mathbb{N}$,

$$D^k = PA^kP^{-1}, \quad \text{ou seja,} \quad A^k = P^{-1}D^kP.$$

Exemplo 32. Uma matriz com valores próprios distintos.

$$A = \begin{bmatrix} 1 & 5 & -1 \\ 0 & -2 & 1 \\ -4 & 0 & 3 \end{bmatrix}.$$

O polinómio característico é dado por

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 1-\lambda & 5 & -1 \\ 0 & -2-\lambda & 1 \\ -4 & 0 & 3-\lambda \end{vmatrix} = \\ &= (1-\lambda)(-2-\lambda)(3-\lambda) - 20 + 4(2+\lambda) = \\ &= (1-\lambda)(-2-\lambda)(3-\lambda) + 4\lambda - 12 = \\ &= (3-\lambda)\left[(\lambda-1)(\lambda+2) - 4\right] = \\ &= (3-\lambda)\left(\lambda^2 + \lambda - 6\right) = \\ &= (3-\lambda)(\lambda-2)(\lambda+3). \end{aligned}$$

Os valores próprios de A são os valores de λ para os quais $\det(A - \lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 3, \quad \lambda_2 = 2 \quad \text{e} \quad \lambda_3 = -3.$$

Os vectores próprios de A associados ao valor próprio λ são os vectores não nulos $v \in \mathbb{R}^3$ para os quais

$$(A - \lambda I)\, v = 0,$$

isto é, são os vectores não nulos de $\mathcal{N}\,(A - \lambda I)$.

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$. Tem-se

$$\mathcal{N}\,(A - \lambda_1 I) = \mathcal{N}\left(\begin{bmatrix} -2 & 5 & -1 \\ 0 & -5 & 1 \\ -4 & 0 & 0 \end{bmatrix}\right) = L\left(\{(0,1,5)\}\right).$$

Logo, o subespaço próprio E_{λ_1} é dado por

$$E_{\lambda_1} = \mathcal{N}\,(A - \lambda_1 I) = L\left(\{(0,1,5)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$ são

$$v = (0, s, 5s)\,, \ \text{ com } \ s \in \mathbb{R} \backslash \{0\}\,.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$. Tem-se

$$\mathcal{N}\,(A - \lambda_2 I) = \mathcal{N}\left(\begin{bmatrix} -1 & 5 & -1 \\ 0 & -4 & 1 \\ -4 & 0 & 1 \end{bmatrix}\right) = L\left(\{(1,1,4)\}\right).$$

Logo, o subespaço próprio E_{λ_2} é dado por

$$E_{\lambda_2} = \mathcal{N}\,(A - \lambda_2 I) = L\left(\{(1,1,4)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$ são

$$v = (s, s, 4s)\,, \ \text{ com } \ s \in \mathbb{R} \backslash \{0\}\,.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_3 = -3$. Tem-se

$$\mathcal{N}\,(A - \lambda_3 I) = \mathcal{N}\left(\begin{bmatrix} 4 & 5 & -1 \\ 0 & 1 & 1 \\ -4 & 0 & 6 \end{bmatrix}\right) = L\left(\{(3,-2,2)\}\right).$$

Logo, o subespaço próprio E_{λ_3} é dado por

$$E_{\lambda_3} = \mathcal{N}\,(A - \lambda_3 I) = L\left(\{(3,-2,2)\}\right).$$

Os vectores próprios de A associados ao valor próprio $\lambda_3 = -3$ são

$$v = (3s, -2s, 2s)\,, \ \text{ com } \ s \in \mathbb{R} \backslash \{0\}\,.$$

Atendendo a que os valores próprios de A são distintos, os vectores próprios de A associados a esses valores próprios são linearmente independentes. Como $\dim \mathbb{R}^3 = 3$, então 3 vectores em $\mathbb{R}^3$ linearmente independentes formarão desde logo uma base de $\mathbb{R}^3$. Logo, o conjunto

$$\mathcal{B} = \{(0,1,5),(1,1,4),(3,-2,2)\}$$

é uma base de $\mathbb{R}^3$. Deste modo, temos uma base de $\mathbb{R}^3$ formada só por vectores próprios de A. Logo, a matriz A é diagonalizável, isto é, existe uma matriz invertível P^{-1} diagonalizante tal que a matriz PAP^{-1} é diagonal, tendo-se

$$D = PAP^{-1} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}, \quad \text{com} \quad P^{-1} = \begin{bmatrix} 0 & 1 & 3 \\ 1 & 1 & -2 \\ 5 & 4 & 2 \end{bmatrix}.$$

Note que cada coluna de P^{-1} é formada pelo vector próprio associado ao valor próprio respectivo e na posição respectiva.

Exemplo 33. Uma matriz com valores próprios repetidos mas diagonalizável.

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}.$$

O polinómio característico é dado por

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 2 \\ 3 & 3 & 4-\lambda \end{vmatrix} = \\ &= (2-\lambda)(3-\lambda)(4-\lambda) + 6 + 6 - 3(3-\lambda) - 6(2-\lambda) - 2(4-\lambda) = \\ &= -\lambda^3 + 9\lambda^2 - 15\lambda + 7 = \\ &= -(\lambda-1)(\lambda-1)(\lambda-7). \end{aligned}$$

Os valores próprios de A são os valores de λ para os quais $\det(A - \lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 1 \quad \text{e} \quad \lambda_2 = 7.$$

Os vectores próprios de A associados ao valor próprio λ são os vectores não nulos $v \in \mathbb{R}^3$ para os quais

$$(A - \lambda I)\, v = 0,$$

isto é, são os vectores não nulos de $\mathcal{N}(A - \lambda I)$.

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_1 = 1$. Tem-se

$$\mathcal{N}(A - \lambda_1 I) = \mathcal{N}\left(\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}\right) = L(\{(-1,1,0),(-1,0,1)\}).$$

Logo, o subespaço próprio E_{λ_1} é dado por

$$E_{\lambda_1} = \mathcal{N}(A - \lambda_1 I) = L(\{(-1,1,0),(-1,0,1)\}).$$

Os vectores próprios de A associados ao valor próprio $\lambda_1 = 1$ são

$$v = (-s-t, s, t), \text{ com } s \neq 0 \text{ ou } t \neq 0.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_2 = 7$. Tem-se

$$\mathcal{N}(A - \lambda_2 I) = \mathcal{N}\left(\begin{bmatrix} -5 & 1 & 1 \\ 2 & -4 & 2 \\ 3 & 3 & -3 \end{bmatrix}\right) = L\left(\{(1,2,3)\}\right).$$

Logo, o subespaço próprio E_{λ_2} é dado por: $E_{\lambda_2} = \mathcal{N}(A - \lambda_2 I) = L(\{(1,2,3)\})$. Os vectores próprios de A associados ao valor próprio $\lambda_2 = 7$ são

$$v = (s, 2s, 3s), \quad \text{com} \quad s \in \mathbb{R} \setminus \{0\}.$$

Atendendo a que $\dim E_{\lambda_1} + \dim E_{\lambda_2} = 3$, podemos ter a seguinte base de $\mathbb{R}^3$ formada só por vectores próprios de A

$$\mathcal{B} = \{(-1,1,0), (-1,0,1), (1,2,3)\}.$$

Logo, a matriz A é diagonalizável, isto é, existe uma matriz invertível P^{-1} diagonalizante tal que a matriz PAP^{-1} é diagonal, tendo-se

$$D = PAP^{-1} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \quad \text{com} \quad P^{-1} = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}.$$

Note que cada coluna de P^{-1} é formada pelo vector próprio associado ao valor próprio respectivo e na posição respectiva.

Exemplo 34. Uma matriz com valores próprios repetidos e não diagonalizável.

$$A = \begin{bmatrix} 7 & 5 & -1 \\ 0 & -2 & 1 \\ 20 & 0 & 3 \end{bmatrix}.$$

O polinómio característico é dado por

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 7-\lambda & 5 & -1 \\ 0 & -2-\lambda & 1 \\ 20 & 0 & 3-\lambda \end{vmatrix} = \\ &= (7-\lambda)(-2-\lambda)(3-\lambda) + 100 - 20(2+\lambda) = \\ &= (3-\lambda)\left[(7-\lambda)(-2-\lambda) + 20\right] = \\ &= (3-\lambda)\left(\lambda^2 - 5\lambda + 6\right) = \\ &= (3-\lambda)(\lambda-3)(\lambda-2). \end{aligned}$$

Os valores próprios de A são os valores de λ para os quais $\det(A - \lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 3 \quad \text{e} \quad \lambda_2 = 2.$$

Os vectores próprios de A associados ao valor próprio λ são os vectores não nulos $v \in \mathbb{R}^3$ para os quais

$$(A - \lambda I)\, v = 0,$$

isto é, são os vectores não nulos de $\mathcal{N}(A-\lambda I)$.

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$. Tem-se

$$\mathcal{N}(A-\lambda_1 I) = \mathcal{N}\left(\begin{bmatrix} 4 & 5 & -1 \\ 0 & -5 & 1 \\ 20 & 0 & 0 \end{bmatrix}\right) = L(\{(0,1,5)\}).$$

Logo, o subespaço próprio E_{λ_1} é dado por: $E_{\lambda_1} = \mathcal{N}(A-\lambda_1 I) = L(\{(0,1,5)\})$. Os vectores próprios de A associados ao valor próprio $\lambda_1 = 3$ são

$$v = (0, s, 5s), \text{ com } s \in \mathbb{R}\setminus\{0\}.$$

Determinemos os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$. Tem-se

$$\mathcal{N}(A-\lambda_2 I) = \mathcal{N}\left(\begin{bmatrix} 5 & 5 & -1 \\ 0 & -4 & 1 \\ 20 & 0 & 1 \end{bmatrix}\right) = L(\{(1,-5,-20)\}).$$

Logo, o subespaço próprio E_{λ_2} é dado por: $E_{\lambda_2} = \mathcal{N}(A-\lambda_2 I) = L(\{(1,-5,-20)\})$. Os vectores próprios de A associados ao valor próprio $\lambda_2 = 2$ são

$$v = (s, -5s, -20s), \text{ com } s \in \mathbb{R}\setminus\{0\}.$$

Atendendo a que

$$\dim E_{\lambda_1} + \dim E_{\lambda_2} = 2 < 3,$$

não é possível ter uma base de $\mathbb{R}^3$ formada só por vectores próprios de A. Logo, a matriz A não é diagonalizável, isto é, não existe uma matriz invertível P^{-1} diagonalizante tal que a matriz PAP^{-1} seja diagonal.

Exemplo 35. Uma matriz com apenas um valor próprio real.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

O polinómio característico é dado por

$$\det(A-\lambda I) = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{vmatrix} = \lambda^2(1-\lambda) + (1-\lambda) = (1-\lambda)\left(\lambda^2+1\right).$$

Os valores próprios de A são os valores de λ para os quais $\det(A-\lambda I) = 0$. Logo, os valores próprios de A são

$$\lambda_1 = 1, \quad \lambda_2 = i \quad \text{e} \quad \lambda_3 = -i.$$

Logo, a matriz A não é diagonalizável numa matriz de entradas reais, isto é, não existe uma matriz invertível P^{-1} diagonalizante tal que a matriz PAP^{-1} seja diagonal com entradas reais. No entanto e atendendo a que os três valores próprios são distintos, a matriz A é diagonalizável numa matriz de entradas complexas:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix}$$

Exemplo 36. A sucessão de Fibonacci (Leonardo de Pisa, 1202). Seja $(v_n)_{n\in\mathbb{N}}$ tal que

$$v_1 = 1, \quad v_2 = 1$$

e

$$v_{n+2} = v_n + v_{n+1}, \;\; n \in \mathbb{N}.$$

Considerando a igualdade $v_{n+1} = v_{n+1}$, podemos escrever o sistema

$$\begin{cases} v_{n+1} = v_{n+1} \\ v_{n+2} = v_n + v_{n+1} \end{cases} \qquad \text{isto é} \qquad \begin{bmatrix} v_{n+1} \\ v_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_n \\ v_{n+1} \end{bmatrix}$$

para todo o $n \in \mathbb{N}$. Aplicando sucessivamente a igualdade anterior tem-se

$$\begin{bmatrix} v_{n+1} \\ v_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_n \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_{n-1} \\ v_n \end{bmatrix} =$$
$$= \cdots = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Calculemos agora os valores próprios de $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$:

$$\det \begin{bmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} = 0 \Leftrightarrow \lambda^2 - \lambda - 1 = 0 \Leftrightarrow \left(\lambda = \frac{1+\sqrt{5}}{2} \quad \text{ou} \quad \lambda = \frac{1-\sqrt{5}}{2} \right).$$

Valores próprios: $\lambda_1 = \frac{1+\sqrt{5}}{2}$ e $\lambda_2 = \frac{1-\sqrt{5}}{2}$.

Atendendo a que

$$\mathcal{N} \begin{bmatrix} -\lambda_1 & 1 \\ 1 & 1-\lambda_1 \end{bmatrix} = \mathcal{N} \begin{bmatrix} 0 & 1+\lambda_1-\lambda_1^2 \\ 1 & 1-\lambda_1 \end{bmatrix} =$$
$$= \mathcal{N} \begin{bmatrix} 0 & 0 \\ 1 & \frac{1-\sqrt{5}}{2} \end{bmatrix} = L\left(\left\{ \left(\frac{-1+\sqrt{5}}{2}, 1 \right) \right\} \right)$$

$\left(\frac{-1+\sqrt{5}}{2}, 1\right)$ é um vector próprio associado ao valor próprio $\frac{1+\sqrt{5}}{2}$, sendo todos os vectores próprios associados ao valor próprio $\frac{1+\sqrt{5}}{2}$ dados por

$$L\left(\left\{ \left(\frac{-1+\sqrt{5}}{2}, 1 \right) \right\} \right) \setminus \{(0,0)\}.$$

Atendendo a que

$$\mathcal{N} \begin{bmatrix} -\lambda_2 & 1 \\ 1 & 1-\lambda_2 \end{bmatrix} = \mathcal{N} \begin{bmatrix} 0 & 1+\lambda_2-\lambda_2^2 \\ 1 & 1-\lambda_2 \end{bmatrix} =$$

$$= \mathcal{N}\begin{bmatrix} 0 & 0 \\ 1 & \frac{1+\sqrt{5}}{2} \end{bmatrix} = L\left(\left\{\left(-\frac{1+\sqrt{5}}{2}, 1\right)\right\}\right)$$

$\left(-\frac{1+\sqrt{5}}{2}, 1\right)$ é um vector próprio associado ao valor próprio $\frac{1-\sqrt{5}}{2}$, sendo todos os vectores próprios associados ao valor próprio $\frac{1-\sqrt{5}}{2}$ dados por

$$L\left(\left\{\left(-\frac{1+\sqrt{5}}{2}, 1\right)\right\}\right) \setminus \{(0,0)\}.$$

Como existe uma base de $\mathbb{R}^2$ formada só por vectores próprios (os dois valores próprios são distintos logo os vectores próprios correspondentes são linearmente independentes) então a matriz $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ é diagonalizável. Assim, fazendo

$$P^{-1} = \begin{bmatrix} \frac{-1+\sqrt{5}}{2} & -\frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \qquad \text{tem-se} \qquad P = \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{5+\sqrt{5}}{10} \\ -\frac{\sqrt{5}}{5} & \frac{5-\sqrt{5}}{10} \end{bmatrix}$$

e

$$D = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} = P\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}P^{-1} \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = P^{-1}\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}P.$$

Logo

$$\begin{bmatrix} v_{n+1} \\ v_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left(P^{-1}\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}P\right)^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} =$$

$$= P^{-1}\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}^n P\begin{bmatrix} 1 \\ 1 \end{bmatrix} = P^{-1}\begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix}P\begin{bmatrix} 1 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{-1+\sqrt{5}}{2} & -\frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix}\begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{5+\sqrt{5}}{10} \\ -\frac{\sqrt{5}}{5} & \frac{5-\sqrt{5}}{10} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{5-\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n + \frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n \\ \frac{5-3\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n + \frac{3\sqrt{5}+5}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n \end{bmatrix}.$$

Isto é,

$$v_{n+1} = \frac{5-\sqrt{5}}{10}\left(\frac{1-\sqrt{5}}{2}\right)^n + \frac{5+\sqrt{5}}{10}\left(\frac{1+\sqrt{5}}{2}\right)^n$$

para todo o $n \in \mathbb{N}$, com $v_1 = 1$.

Verifique que (por exemplo) $v_2 = 1$, $v_3 = 2$, $v_4 = 3$.

Exemplo 37. (Um processo de difusão.) Considere duas células adjacentes separadas por uma membrana permeável e suponha que um fluído passa da 1^a célula para a 2^a a uma taxa (em mililitros por minuto) numericamente igual a 4 vezes o volume (em mililitros) do fluído da 1^a célula. Em seguida, passa da 2^a célula para a 1^a a uma taxa (em mililitros por minuto) numericamente igual a 5 vezes o volume (em mililitros) do fluído da 2^a célula.

Sejam $v_1(t)$ e $v_2(t)$ respectivamente o volume da 1^a célula e o volume da 2^a célula no instante t. Suponha que inicialmente a primeira célula tem 10 mililitros de fluído e que a segunda tem 8 mililitros de fluído, isto é $v_1(0) = 10$ e $v_2(0) = 8$.

Determinemos o volume de fluído de cada célula no instante t.

Tem-se

$$\begin{cases} v_1'(t) = -4v_1(t) \\ \\ v_2'(t) = 4v_1(t) - 5v_2(t) \end{cases}$$

isto é

$$\begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}. \quad (*)$$

-4 e -5 são os valores próprios da matriz $\begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix}$, sendo os vectores próprios associados $(1,4)$ e $(0,1)$ respectivamente.

Como existe uma base de $\mathbb{R}^2$ formada só por vectores próprios (os dois valores próprios são distintos logo os vectores próprios correspondentes são linearmente independentes) então a matriz $\begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix}$ é diagonalizável. Assim, fazendo

$$P^{-1} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \quad \text{tem-se} \quad P = \begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix}$$

e

$$\Leftrightarrow D = \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} = P \begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix} P^{-1} \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} -4 & 0 \\ 4 & -5 \end{bmatrix} = P^{-1} \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} P.$$

o sistema (*) é equivalente a

$$\begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix} = \left(P^{-1} \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} P \right) \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow P \begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix} \left(P \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \right).$$

Assim, considerando a mudança de variável

$$\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = P \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$$

$$P\begin{bmatrix} v_1'(t) \\ v_2'(t) \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix}\left(P\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}\right) \Leftrightarrow$$

$$\Leftrightarrow \begin{bmatrix} u_1'(t) \\ u_2'(t) \end{bmatrix} = \begin{bmatrix} -4 & 0 \\ 0 & -5 \end{bmatrix}\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} u_1'(t) = -4u_1(t) \\ \\ u_2'(t) = -5u_2(t) \end{cases} \underset{\substack{u_1(t)\neq 0 \\ \text{e } u_2(t)\neq 0,\ \forall t}}{\Leftrightarrow} \begin{cases} \dfrac{u_1'(t)}{u_1(t)} = -4 \\ \\ \dfrac{u_2'(t)}{u_2(t)} = -5 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} \log|u_1(t)| = -4t + k_1 \\ \\ \log|u_2(t)| = -5t + k_2 \end{cases} \Leftrightarrow \begin{cases} u_1(t) = c_1 e^{-4t} \\ \\ u_2(t) = c_2 e^{-5t} \end{cases}$$

com $c_1, c_2 \in \mathbb{R}$. De facto, se $u(t)$ fôr solução de $u'(t) = \alpha u(t)$ então $u(t)\,e^{-\alpha t} = c$ (constante) uma vez que $\left(u(t)\,e^{-\alpha t}\right)' = 0$. Logo $u(t) = ce^{\alpha t}$.

Assim

$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = P^{-1}\begin{bmatrix} c_1 e^{-4t} \\ c_2 e^{-5t} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} c_1 e^{-4t} \\ c_2 e^{-5t} \end{bmatrix} = \begin{bmatrix} c_1 e^{-4t} \\ 4c_1 e^{-4t} + c_2 e^{-5t} \end{bmatrix}.$$

Como

$$\begin{cases} v_1(0) = 10 \\ \\ v_2(0) = 8 \end{cases}$$

então $c_1 = 10$ e $c_2 = -32$ e assim a solução geral do sistema de equações diferenciais lineares

$$\begin{cases} v_1'(t) = -4v_1(t) \\ \\ v_2'(t) = 4v_1(t) - 5v_2(t) \end{cases}$$

com os valores iniciais

$$\begin{cases} v_1(0) = 10 \\ \\ v_2(0) = 8 \end{cases}$$

é dada por

$$\begin{bmatrix} v_1(t) \\ \\ v_2(t) \end{bmatrix} = \begin{bmatrix} 10e^{-4t} \\ \\ 40e^{-4t} - 32e^{-5t} \end{bmatrix} = 10\begin{bmatrix} 1 \\ 4 \end{bmatrix} e^{-4t} - 32\begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-5t}.$$

Produtos internos e ortogonalização

Definição 49. Sejam V um espaço linear real e $\mathbf{0}$ o vector nulo de V. Chama-se **produto interno** em V a uma aplicação

$$\langle , \rangle : V \times V \to \mathbb{R}$$

$$(u, v) \;\; \to \langle u, v \rangle$$

que verifique as três condições seguintes.

(i) Simetria: para todos os $u, v \in V$

$$\langle u, v \rangle = \langle v, u \rangle .$$

(ii) Linearidade: para todo o $v \in V$ (fixo)

$$\langle \alpha u + \beta w, v \rangle = \alpha \langle u, v \rangle + \beta \langle w, v \rangle$$

para todos os $u, w \in V$ e $\alpha, \beta \in \mathbb{R}$. Isto é, para todo o $v \in V$ (fixo) a aplicação

$$V \to \mathbb{R}$$

$$u \to \langle u, v \rangle$$

é linear.

(iii) Positividade: para todo o $u \in V$ tal que $u \neq \mathbf{0}$,

$$\langle u, u \rangle > 0.$$

Tendo-se $\langle u, u \rangle = 0$ se e só se $u = \mathbf{0}$.

Observação 25. **(a)** Um produto interno num espaço linear real é uma forma **bilinear**, **simétrica** e **definida positiva**.

(b) Num espaço linear V sobre $\mathbb{C}$ (espaço linear complexo), um produto interno é uma aplicação que a cada par de vectores $(u, v) \in V \times V$ associa o número complexo $\langle u, v \rangle$ e que verifica as seguintes condições:

(i) Para todos os $u, v \in V$

$$\langle u, v \rangle = \overline{\langle v, u \rangle}.$$

(ii) Para todo o $v \in V$ (fixo) tem-se

$$\langle \alpha u + \beta w, v \rangle = \overline{\alpha} \langle u, v \rangle + \overline{\beta} \langle w, v \rangle$$

para todos os $u, w \in V$ e $\alpha, \beta \in \mathbb{C}$, (onde por exemplo $\overline{\alpha} = a - bi$ se $\alpha = a + bi$).

(iii) Para todo o $u \in V$ tal que $u \neq \mathbf{0}$,

$$\langle u, u \rangle > 0.$$

Tendo-se $\langle u, u \rangle = 0$ se e só se $u = \mathbf{0}$.

(c) A um espaço linear real de dimensão finita com um produto interno chama-se **espaço euclidiano**. A um espaço linear complexo de dimensão finita com um produto interno chama-se **espaço unitário**.

Observação 26. **(i)** Seja V um espaço euclidiano. Seja $\mathcal{B} = \{w_1, w_2, ..., w_n\}$ uma base ordenada de V. Sejam $u, v \in V$. Sejam

$$\alpha_1, \alpha_2, ..., \alpha_n \quad \text{e} \quad \beta_1, \beta_2, ..., \beta_n$$

as coordenadas de u e de v na base ordenada $\mathcal{B}$ respectivamente, isto é,

$$u = \sum_{i=1}^{n} \alpha_i w_i \quad \text{e} \quad v = \sum_{i=1}^{n} \beta_i w_i.$$

Logo,

$$\langle u, v \rangle = \left\langle \sum_{i=1}^{n} \alpha_i w_i, \sum_{i=1}^{n} \beta_i w_i \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \langle w_i, w_j \rangle =$$

$$= \begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{bmatrix} \begin{bmatrix} \langle w_1, w_1 \rangle & \langle w_1, w_2 \rangle & \ldots & \langle w_1, w_n \rangle \\ \langle w_2, w_1 \rangle & \langle w_2, w_2 \rangle & \ldots & \langle w_2, w_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle w_n, w_1 \rangle & \langle w_n, w_2 \rangle & \ldots & \langle w_n, w_n \rangle \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = ([u]_{\mathcal{B}})^T G_{\mathcal{B}} [v]_{\mathcal{B}}.$$

Assim, fixando uma base ordenada $\mathcal{B} = \{w_1, w_2, ..., w_n\}$ de V, a aplicação $\langle , \rangle : V \times V \to \mathbb{R}$ que a cada $(u, v) \in V \times V$ faz corresponder $\langle u, v \rangle$, é um produto interno em V se e só se a matriz

$$G_{\mathcal{B}} = \begin{bmatrix} \langle w_1, w_1 \rangle & \langle w_1, w_2 \rangle & \ldots & \langle w_1, w_n \rangle \\ \langle w_2, w_1 \rangle & \langle w_2, w_2 \rangle & \ldots & \langle w_2, w_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle w_n, w_1 \rangle & \langle w_n, w_2 \rangle & \ldots & \langle w_n, w_n \rangle \end{bmatrix}$$

fôr simétrica ($G_{\mathcal{B}} = G_{\mathcal{B}}^T$) e definida positiva ($([u]_{\mathcal{B}})^T G_{\mathcal{B}} [u]_{\mathcal{B}} > 0$, para todo o $u \neq \mathbf{0}$). Note-se que atendendo às propriedades referentes às operações matriciais envolvidas, a igualdade

$$\langle u, v \rangle = ([u]_{\mathcal{B}})^T G_{\mathcal{B}} [v]_{\mathcal{B}}$$

equivale à bilinearidade da aplicação $\langle , \rangle : V \times V \to \mathbb{R}$.

(ii) À matriz $G_{\mathcal{B}}$ anterior dá-se o nome de matriz de Gram ou matriz da métrica do produto interno.

(iii) Num próximo capítulo, como consequência da diagonalização ortogonal, sendo $G_{\mathcal{B}}$ simétrica ($G_{\mathcal{B}} = G_{\mathcal{B}}^T$), será estabelecida a equivalência:

$$\left(([u]_{\mathcal{B}})^T G_{\mathcal{B}} [u]_{\mathcal{B}} > 0, \quad \text{para todo o} \quad u \neq \mathbf{0}\right) \Leftrightarrow (\text{todos os valores próprios de } G_{\mathcal{B}} \text{ são positivos}).$$

(iv) Observe-se ainda que no caso de se ter um espaço unitário, a matriz $G_{\mathcal{B}}$ tem os valores próprios (num próximo capítulo) todos positivos e é **hermitiana**, isto é, é tal que $G_{\mathcal{B}} = \overline{G_{\mathcal{B}}}^T$, (onde $\overline{G_{\mathcal{B}}}$ é a matriz que se obtem de $G_{\mathcal{B}}$ passando todas as entradas desta ao complexo conjugado), tendo-se

$$\langle u, v \rangle = \begin{bmatrix} \overline{\alpha_1} & \overline{\alpha_2} & \ldots & \overline{\alpha_n} \end{bmatrix} G_{\mathcal{B}} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}.$$

Teorema 45. Seja V um espaço euclidiano com $\dim V = n$. Seja $\{w_1, w_2, ..., w_n\}$ uma base ordenada de V. Então, uma aplicação

$$\langle , \rangle : V \times V \to \mathbb{R}$$

é um produto interno (em V) se e só se

$$\langle u, v \rangle = \begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{bmatrix} G_{\mathcal{B}} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = ([u]_{\mathcal{B}})^T G_{\mathcal{B}} [v]_{\mathcal{B}},$$

com

$$u = \alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_n w_n \qquad v = \beta_1 w_1 + \beta_2 w_2 + ... + \beta_n w_n.$$

e $G_{\mathcal{B}}$ é uma matriz simétrica cujos valores próprios são todos positivos, dada por:

$$G_{\mathcal{B}} = \begin{bmatrix} \langle w_1, w_1 \rangle & \langle w_1, w_2 \rangle & \ldots & \langle w_1, w_n \rangle \\ \langle w_2, w_1 \rangle & \langle w_2, w_2 \rangle & \ldots & \langle w_2, w_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle w_n, w_1 \rangle & \langle w_n, w_2 \rangle & \ldots & \langle w_n, w_n \rangle \end{bmatrix}.$$

Exemplo 38. (i) Seja $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por:

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = \alpha_1 \beta_1 + \alpha_2 \beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$. Esta aplicação é um produto interno em $\mathbb{R}^2$ a que se dá o nome de produto interno usual em $\mathbb{R}^2$, uma vez que

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = \alpha_1 \beta_1 + \alpha_2 \beta_2 = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} G_{\mathcal{B}_c} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

com
$$G_{\mathcal{B}_c} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$
A matriz $G_{\mathcal{B}_c}$ é simétrica e o único valor próprio de $G_{\mathcal{B}_c}$ é $1 > 0$.

(ii) Seja $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por:
$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = -2\alpha_1\beta_1 + 3\alpha_2\beta_2,$$
com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$. Esta aplicação não é um produto interno em $\mathbb{R}^2$, uma vez que
$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle = -2\alpha_1\beta_1 + 3\alpha_2\beta_2 = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} G \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad \text{com} \quad G = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}.$$
A matriz G é simétrica, no entanto, os valores próprios de G: -2 e 3 não são ambos positivos.

(iii) O produto interno usual em $\mathbb{R}^n$ é dado por:
$$\langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$
$$(u, v) \to \langle u, v \rangle = u^T v,$$
onde $u^T = \begin{bmatrix} u_1 & u_2 & \ldots & u_n \end{bmatrix}$ e $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$.

(iv) O produto interno usual em $\mathbb{C}^n$ é dado por:
$$\langle , \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$$
$$(u, v) \to \langle u, v \rangle = u^H v,$$
onde $u^H = \overline{u}^T = \begin{bmatrix} \overline{u_1} & \overline{u_2} & \ldots & \overline{u_n} \end{bmatrix}$ e $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$.

(v) Um produto interno em $\mathcal{M}_{m \times n}(\mathbb{R})$.
$$\langle , \rangle : M_{m \times n}(\mathbb{R}) \times M_{m \times n}(\mathbb{R}) \to \mathbb{R}$$
$$(A, B) \to \langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij} = \operatorname{tr}\left(A^T B\right).$$

(vi) Um produto interno em $C([a, b])$.
$$\langle , \rangle : C([a, b]) \times C([a, b]) \to \mathbb{R}$$
$$(f, g) \to \langle f, g \rangle = \int_a^b f(x)\, g(x)\, dx.$$

Prova da positividade: $\langle f, f\rangle > 0$ para toda a função não nula. Seja $f \in C([a,b])$. Seja $x_0 \in [a,b]$ tal que $f(x_0) \neq 0$. Como f^2 é contínua em $[a,b]$, existe um intervalo $I \subset [a,b]$ tal que para todo o $x \in I$

$$(f(x))^2 \geq \frac{(f(x_0))^2}{2}.$$

Logo

$$\langle f, f\rangle = \int_a^b (f(x))^2\, dx \geq \int_I (f(x))^2\, dx \geq \int_I \frac{(f(x_0))^2}{2} dx = \frac{(f(x_0))^2}{2} \int_I dx = \frac{(f(x_0))^2}{2} |I| > 0$$

onde $|I|$ denota o comprimento do intervalo I.

Exemplo 39. $\mathbb{R}^2$ com um produto interno não usual. Seja $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por:

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = 2\alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_2\beta_1 + 3\alpha_2\beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$.

É fácil ver que esta aplicação é simétrica e linear em relação a (α_1, α_2) (fixando (β_1, β_2)). Vejamos por exemplo que a condição

$$\langle (\alpha_1, \alpha_2), (\alpha_1, \alpha_2)\rangle > 0, \quad \text{para todo o } (\alpha_1, \alpha_2) \neq (0,0),$$

é satisfeita.

Atendendo a que

$$\langle (\alpha_1, \alpha_2), (\alpha_1, \alpha_2)\rangle = 2\alpha_1^2 + 2\alpha_1\alpha_2 + 3\alpha_2^2 = \alpha_1^2 + (\alpha_1 + \alpha_2)^2 + 2\alpha_2^2,$$

tem-se

$$\begin{gathered}\langle (\alpha_1, \alpha_2), (\alpha_1, \alpha_2)\rangle = 0 \Leftrightarrow \\ \Leftrightarrow (\alpha_1 = 0 \text{ e } \alpha_1 + \alpha_2 = 0 \text{ e } \alpha_2 = 0) \\ \Leftrightarrow (\alpha_1 = 0 \text{ e } \alpha_2 = 0) \Leftrightarrow (\alpha_1, \alpha_2) = (0,0).\end{gathered}$$

Em alternativa, podemos escrever

$$\begin{aligned}&\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = 2\alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_2\beta_1 + 3\alpha_2\beta_2 = \\ &= \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} G_{\mathcal{B}_c} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \qquad \text{com} \qquad G_{\mathcal{B}_c} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}.\end{aligned}$$

A matriz $G_{\mathcal{B}_c}$ é simétrica e os valores próprios de $G_{\mathcal{B}_c}$: $\frac{5+\sqrt{5}}{2}$ e $\frac{5-\sqrt{5}}{2}$ são ambos positivos.

Definição 50. Sejam V um espaço linear com um produto interno e $\mathbf{0}$ o vector nulo de V. Sejam $u, v \in V$.

(i) Chama-se **norma** de u a:

$$\|u\| = \sqrt{\langle u, u\rangle}.$$

(ii) Diz-se que u e v são **ortogonais** se $\langle u, v\rangle = 0$.

(iii) A **projecção ortogonal** de v sobre $u \neq \mathbf{0}$:

$$\operatorname{proj}_u v$$

deverá ser tal que

$$\operatorname{proj}_u v = ku \quad \text{(para um certo escalar } k) \quad \text{e} \quad \langle v - ku, ku\rangle = 0.$$

Como

$$\langle v - ku, u\rangle = 0 \Leftrightarrow \langle v, u\rangle - \overline{k}\,\langle u, u\rangle = 0 \Leftrightarrow k = \frac{\langle u, v\rangle}{\|u\|^2},$$

então a **projecção ortogonal** de v sobre $u \neq \mathbf{0}$ será definida por:

$$\operatorname{proj}_u v = \frac{\langle u, v\rangle}{\|u\|^2} u.$$

No caso de V ser um espaço linear real pode escrever-se: $\operatorname{proj}_u v = \dfrac{\langle v, u\rangle}{\|u\|^2} u$.

(iv) Chama-se **ângulo** entre dois vectores não nulos u e v tais que $\langle u, v\rangle \in \mathbb{R}$ a:

$$\theta = \arccos \frac{\langle u, v\rangle}{\|u\|\,\|v\|}.$$

Note que este ângulo está bem definido atendendo ao próximo teorema.

Observação 27. **(i)** O ângulo θ entre dois vectores não nulos u e v é $\frac{\pi}{2}$ se e só se u e v são ortogonais.

(ii) Para cada $u \in V$ (fixo) com $u \neq \mathbf{0}$, a aplicação $\operatorname{proj}_u : V \to V$ que a cada $v \in V$ faz corresponder $\operatorname{proj}_u v$ satisfaz:

$$\operatorname{proj}_u (\alpha v + \beta w) = \alpha \operatorname{proj}_u (v) + \beta \operatorname{proj}_u (w),$$

para todos os $v, w \in V$ e $\alpha, \beta \in \mathbb{R}$. A uma aplicação que verifique a condição anterior chamaremos transformação linear.

(iii) Sejam $u, v \in V \setminus \{\mathbf{0}\}$. Tem-se

$$\operatorname{proj}_u v = \frac{\langle u, v\rangle}{\|u\|^2} u = \|v\| \frac{\langle u, v\rangle}{\|u\|\,\|v\|} \frac{1}{\|u\|} u = (\|v\| \cos\theta) \frac{1}{\|u\|} u.$$

Teorema 46. **(i) Desigualdade de Cauchy-Schwarz.** Seja V um espaço linear com um produto interno. Então, para todos os $u, v \in V$,

$$|\langle u, v\rangle| \leq \|u\|\,\|v\|$$

(ii) Sejam $u, v \in V$. Tem-se:

$$|\langle u, v\rangle| = \|u\| \, \|v\| \Leftrightarrow \{u, v\} \text{ é linearmente dependente.}$$

Dem. **(i)** Sejam $u, v \in V$. Se $u = \mathbf{0}$ a desigualdade é satisfeita. Se $u \neq \mathbf{0}$, tem-se

$$0 \leq \|v - \operatorname{proj}_u v\|^2 = \left\| v - \frac{\langle u, v\rangle}{\|u\|^2} u \right\|^2 = \left\langle v - \frac{\langle u, v\rangle}{\|u\|^2} u, v - \frac{\langle u, v\rangle}{\|u\|^2} u \right\rangle =$$

$$= \|v\|^2 + \frac{\langle u, v\rangle}{\|u\|^2} \frac{\overline{\langle u, v\rangle}}{\|u\|^2} \|u\|^2 - 2 \frac{\overline{\langle u, v\rangle}}{\|u\|^2} |\langle u, v\rangle| = \|v\|^2 - \frac{|\langle u, v\rangle|^2}{\|u\|^2}$$

e assim

$$|\langle u, v\rangle| \leq \|u\| \, \|v\| \, .$$

(ii) Se $u = \mathbf{0}$ então:

$$|\langle u, v\rangle| = \|u\| \, \|v\| \Leftrightarrow \{u, v\} \text{ é linearmente dependente.}$$

Se $u \neq \mathbf{0}$ então:

$$\{u, v\} \text{ é linearmente dependente} \Leftrightarrow 0 = \|v - \operatorname{proj}_u v\|^2 = \|v\|^2 - \frac{|\langle u, v\rangle|^2}{\|u\|^2} \Leftrightarrow$$

$$\Leftrightarrow |\langle u, v\rangle| = \|u\| \, \|v\| \, .$$

Teorema 47. Teorema de Pitágoras. Seja V um espaço linear real com um produto interno. Sejam $u, v \in V$. Tem-se u e v ortogonais se e só se

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 \, .$$

Dem.

$$\|u - v\|^2 = \langle u - v, u - v\rangle = \langle u, u\rangle - \langle v, u\rangle - \langle u, v\rangle + \langle v, v\rangle = \|u\|^2 - 2 \langle u, v\rangle + \|v\|^2 = \|u\|^2 + \|v\|^2$$

se e só se

$$\langle u, v\rangle = 0,$$

isto é, se e só se u e v forem ortogonais.

Observação 28. **(i)** Num espaço euclidiano, o teorema de Pitágoras pode ser enunciado do seguinte modo:

$$\|v\|^2 = \|\operatorname{proj}_u v\|^2 + \|v - \operatorname{proj}_u v\|^2$$

para todos os u, v.

$$\|v\|^2 = \|\operatorname{proj}_u v\|^2 + \|v - \operatorname{proj}_u v\|^2 \geq \|\operatorname{proj}_u v\|^2$$

(ii) Num espaço euclidiano, a desigualdade de Cauchy-Schwarz poderia ter sido provada recorrendo ao teorema de Pitágoras, uma vez que

$$\|\operatorname{proj}_u v\|^2 = \|v\|^2 - \|v - \operatorname{proj}_u v\|^2 \leq \|v\|^2 \Leftrightarrow$$

$$\Leftrightarrow \left(\frac{|\langle v, u\rangle|}{\|u\|^2}\right)^2 \|u\|^2 \leq \|v\|^2 \Leftrightarrow |\langle u, v\rangle| \leq \|u\| \|v\| .$$

(iii) Em $\mathbb{R}^2$ com o produto interno usual, a desigualdade de Cauchy-Schwarz é dada por

$$|\alpha_1\beta_1 + \alpha_2\beta_2| \leq \sqrt{\alpha_1^2 + \alpha_2^2}\sqrt{\beta_1^2 + \beta_2^2},$$

uma vez que

$$\langle(\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = \alpha_1\beta_1 + \alpha_2\beta_2,$$

com $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \in \mathbb{R}^2$.

(iv) Em $\mathbb{R}^n$ com o produto interno usual, a desigualdade de Cauchy-Schwarz é dada por

$$\left|\sum_{i=1}^{n} \alpha_i\beta_i\right| \leq \sqrt{\sum_{i=1}^{n} \alpha_i^2}\sqrt{\sum_{i=1}^{n} \beta_i^2},$$

uma vez que

$$\langle(\alpha_1, ..., \alpha_n), (\beta_1, ..., \beta_n)\rangle = \alpha_1\beta_1 + ... + \alpha_n\beta_n,$$

com $(\alpha_1, ..., \alpha_n), (\beta_1, ..., \beta_n) \in \mathbb{R}^n$.

Teorema 48. Sejam V um espaço linear com um produto interno e $\mathbf{0}$ o vector nulo de V. Sejam $u, v \in V$ e λ escalar. A norma é uma aplicação $\|\| : V \to \mathbb{R}$ que satisfaz as seguintes propriedades.

(i) Positividade: $\|u\| > 0$ se $u \neq \mathbf{0}$.

(ii) Homogeneidade: $\|\lambda u\| = |\lambda| \|u\|$

(iii) Desigualdade triangular: $\|u + v\| \leq \|u\| + \|v\|$

Dem. (iii)

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 + 2\operatorname{Re}\langle u, v\rangle \leq$$

$$\leq \|u\|^2 + \|v\|^2 + 2|\langle u, v\rangle| \underset{\text{C.-S.}}{\leq} \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| =$$

$$= (\|u\| + \|v\|)^2$$

logo

$$\|u + v\| \leq \|u\| + \|v\| .$$

Definição 51. Pode definir-se **norma** num espaço linear V, sem estar associada a qualquer produto interno, como sendo uma aplicação de V em $\mathbb{R}$ que satisfaz as propriedades do teorema anterior. A um espaço linear com uma norma chama-se **espaço normado**.

Observação 29. Seja V um espaço linear real com um produto interno. Sejam $u, v \in V$. Tem-se

$$\langle u, v\rangle = \frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right).$$

Teorema 49. Seja V um espaço normado. Sejam $u, v \in V$. Então, a norma pode dar origem a um produto interno definido por

$$\langle u, v\rangle = \frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right).$$

se e só se

$$\|u-v\|^2 + \|u+v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Esta última equação é conhecida por **lei do paralelogramo**.

Exemplo 40. Uma norma que não dá origem a um produto interno. Seja $\|\| : \mathbb{R}^2 \to \mathbb{R}$ a aplicação definida por

$$\|(\alpha_1, \alpha_2)\| = |\alpha_1| + |\alpha_2|,$$

com $(\alpha_1, \alpha_2) \in \mathbb{R}^2$. É fácil verificar que esta aplicação satisfaz as três condições da norma. Logo, é uma norma. No entanto, é também fácil verificar que esta norma não satisfaz a lei do paralelogramo. Logo, esta norma não poderá originar um produto interno.

Definição 52. Sejam V um espaço linear com um produto interno e $S \subset V$. Diz-se que S é **ortogonal** se para todos os $u, v \in S$ com $u \neq v$, se tiver

$$\langle u, v\rangle = 0.$$

Diz-se que S é **ortonormado** se fôr ortogonal e se, para todo o $u \in S$, se tiver

$$\|u\| = 1.$$

Teorema 50. Sejam V um espaço linear com um produto interno e $S \subset V$. Seja $\mathbf{0}$ o vector nulo de V. Se S é ortogonal e $\mathbf{0} \notin S$ então S é linearmente independente. Em particular, se $n = \dim V$ então qualquer conjunto S ortogonal de n vectores não nulos é uma base de V.

Teorema 51. Seja V um espaço euclidiano com $\dim V = n$. Seja $\mathcal{B} = \{u_1, ..., u_n\}$ uma base (ordenada) ortogonal de V. Então, as coordenadas de um vector $v \in V$ em relação à base (ordenada) $\mathcal{B}$ são dadas por:

$$\alpha_j = \frac{\langle v, u_j \rangle}{\langle u_j, u_j \rangle},$$

com $j = 1, ..., n$. Se $\mathcal{B}$ fôr ortonormada então as coordenadas de um vector $v \in V$ em relação à base (ordenada) $\mathcal{B}$ são dadas por:

$$\alpha_j = \langle v, u_j \rangle,$$

com $j = 1, ..., n$.

Teorema 52. Seja V um espaço euclidiano com $\dim V = n$. Seja $\mathcal{B} = \{w_1, ..., w_n\}$ uma base (ordenada) ortonormada de V. Então, para todos os $u, v \in V$, tem-se

$$\langle u, v \rangle = \sum_{i=1}^{n} \langle u, w_i \rangle \langle v, w_i \rangle \qquad \text{(fórmula de Parseval)}$$

e

$$\|u\| = \sqrt{\sum_{i=1}^{n} \langle u, w_i \rangle^2}.$$

Observação 30. Seja V um espaço euclidiano com $\dim V = n$. Seja $\mathcal{B} = \{w_1, ..., w_n\}$ uma base (ordenada) ortonormada de V. Sejam $u, v \in V$, com

$$u = \alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_n w_n \qquad v = \beta_1 w_1 + \beta_2 w_2 + ... + \beta_n w_n.$$

Então a fórmula de Parseval é dada por:

$$\langle u, v \rangle = \sum_{i=1}^{n} \alpha_i \beta_i = \alpha_1 \beta_1 + \alpha_2 \beta_2 + ... + \alpha_n \beta_n$$

e tem-se

$$\|u\| = \sqrt{\sum_{i=1}^{n} \alpha_i^2}.$$

Notação 3. Sejam V um espaço linear com um produto interno e $\mathbf{0}$ o vector nulo de V. Para qualquer $v \in V$, com $v \neq \mathbf{0}$, o vector $\frac{1}{\|v\|} v$ será denotado por $\frac{v}{\|v\|}$.

Teorema 53. Método de ortogonalização de Gram-Schmidt. Seja V um espaço euclidiano (ou unitário) não nulo. Seja U um subespaço de V. Então U tem bases ortonormadas. Mais concretamente, seja

$$\{v_1, v_2, ..., v_k\}$$

uma base de U e sejam

$$\begin{aligned} u_1 &= v_1, \\ u_2 &= v_2 - \operatorname{proj}_{u_1} v_2, \\ &\cdots \\ u_k &= v_k - \operatorname{proj}_{u_1} v_k - ... - \operatorname{proj}_{u_{k-1}} v_k \end{aligned}$$

então

(i) $L(\{u_1, u_2, ..., u_k\}) = L(\{v_1, v_2, ..., v_k\}) = U$;

(ii) o conjunto $\{u_1, u_2, ..., u_k\}$ é uma base ortogonal de U.

(iii) o conjunto $\left\{\frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}, ..., \frac{u_k}{\|u_k\|}\right\}$ é uma base ortonormada de U.

Exemplo 41. Considere-se $\mathbb{R}^4$ com o produto interno usual. Seja

$$U = L(\{(1,1,-1,-1),(1,2,3,4),(2,1,-6,-7),(1,3,7,9)\}).$$

Determinemos a dimensão de U e uma base ortonormada para U. Tem-se

$$\begin{bmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \\ -1 & 3 & -6 & 7 \\ -1 & 4 & -7 & 9 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 4 & -4 & 8 \\ 0 & 5 & -5 & 10 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Logo, o conjunto $\{v_1, v_2\}$, com $v_1 = (1,1,-1,-1)$ e $v_2 = (1,2,3,4)$, é uma base de U e como tal $\dim U = 2$.

Sejam $u_1 = v_1$ e $u_2 = v_2 - \operatorname{proj}_{u_1} v_2$.

Logo, o conjunto $\{u_1, u_2\}$, com $u_1 = (1,1,-1,-1)$ e

$$u_2 = (1,2,3,4) - \frac{1+2-3-4}{4}(1,1,-1,-1) = (2,3,2,3),$$

é uma base ortogonal de U. Uma base ortonormada para U:

$$\left\{\frac{u_1}{\|u_1\|}, \frac{u_2}{\|u_2\|}\right\} = \left\{\left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right), \left(\frac{\sqrt{26}}{13}, \frac{3\sqrt{26}}{26}, \frac{\sqrt{26}}{13}, \frac{3\sqrt{26}}{26}\right)\right\}$$

Teorema 54. Seja $\mathcal{B} = \{u_1, u_2, ..., u_n\}$ uma base (ordenada) de um espaço euclidiano (ou unitário). A base $\mathcal{B}$ é ortonormada se e só se a matriz da métrica $G_{\mathcal{B}}$ em relação a essa base fôr a matriz identidade. Em $\mathbb{R}^n$ o produto interno usual é aquele (o único) em relação ao qual a base canónica é ortonormada.

Teorema 55. Seja $\{v_1, v_2, ..., v_n\}$ uma base (ordenada) de $\mathbb{R}^n$. Então, existe um único produto interno em $\mathbb{R}^n$ para o qual esta base é ortonormada.

Exemplo 42. Considere em $\mathbb{R}^2$ a base (ordenada) $\mathcal{B} = \{v_1, v_2\}$, com $v_1 = (1, 0)$ e $v_2 = (1, 1)$. Vejamos que existe um e um só produto interno para o qual a base $\mathcal{B}$ é ortonormada.

Seja $\mathcal{B}_c^2 = \{(1, 0), (0, 1)\}$ a base canónica de $\mathbb{R}^2$. Sejam $u, v \in \mathbb{R}^2$, com $u = (\alpha_1, \alpha_2)$ e $v = (\beta_1, \beta_2)$, onde α_1, α_2 e β_1, β_2 são as coordenadas na base $\mathcal{B}_c^2$ de u e v respectivamente. Logo, a aplicação $\langle , \rangle : \mathbb{R}^2 \times \mathbb{R}^2$ definida por

$$\langle u, v\rangle = \langle (\alpha_1, \alpha_2), (\beta_1, \beta_2)\rangle = \left(\begin{bmatrix} \alpha_1 - \alpha_2 \\ \alpha_2 \end{bmatrix}\right)^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} \beta_1 - \beta_2 \\ \beta_2 \end{bmatrix}\right) =$$

$$= \alpha_1\beta_1 - \alpha_1\beta_2 - \alpha_2\beta_1 + 2\alpha_2\beta_2 = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

é um produto interno e é o único para o qual a base $\mathcal{B}$ é ortonormada.

NOTE QUE: $G_{\mathcal{B}_c} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$ (é simétrica e os valores próprios $\frac{3+\sqrt{5}}{2}$ e $\frac{3-\sqrt{5}}{2}$ são ambos positivos) é a matriz da métrica em relação a $\mathcal{B}_c$ e $G'_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (é simétrica e o único valor próprio 1 é positivo) a matriz da métrica em relação a $\mathcal{B}$. É fácil verificar que para este produto interno a base $\mathcal{B}$ é ortonormada:

$$\langle (1, 0), (1, 1)\rangle = 0 \quad \text{e} \quad \langle (1, 0), (1, 0)\rangle = \langle (1, 1), (1, 1)\rangle = 1.$$

Definição 53. Sejam V um espaço linear com produto interno e U um subespaço de V. Diz-se que um elemento de V é **ortogonal a** U se fôr ortogonal a todos os elementos de U. Ao conjunto de todos os elementos ortogonais a U chama-se **complemento ortogonal** de U e designa-se por $U^\perp$,

$$U^\perp = \{v \in V : \langle v, u\rangle = 0 \text{ para todo o } u \in U\}.$$

Teorema 56. Seja V um espaço linear com produto interno. Qualquer que seja o subespaço U de V, também $U^\perp$ é um subespaço de V.

Definição 54. Sendo S um subconjunto de V, não necessariamente um subespaço de V, (também) pode definir-se $S^\perp$:

$$S^\perp = \{v \in V : \langle v, u\rangle = 0 \text{ para todo o } u \in S\}.$$

Observação 31. Apesar de S não ser necessariamente um subespaço de V, $S^{\perp}$ é sempre um subespaço de V, tendo-se

$$S^{\perp} = (L(S))^{\perp}.$$

Teorema 57. Seja V um espaço linear com produto interno.

(i) Seja U um subespaço de V. Tem-se

$$U \cap U^{\perp} = \{\mathbf{0}\}.$$

(ii) Seja S um subconjunto de V. Então

$$S \subset \left(S^{\perp}\right)^{\perp}.$$

No próximo teorema ver-se-á que no caso de se ter $\dim V < \infty$, então

$$L(S) = \left(S^{\perp}\right)^{\perp}$$

ou ainda, sendo U um subespaço de V com $\dim V < \infty$, então

$$U = \left(U^{\perp}\right)^{\perp}.$$

(iii) Sejam S_1, S_2 subconjuntos de V. Então

$$S_1 \subset S_2 \Rightarrow (S_2)^{\perp} \subset (S_1)^{\perp}$$

(iv) Seja U um subespaço de V. Se $\{v_1, ..., v_n\}$ é uma base de U então

$$U^{\perp} = \{v \in V : \langle v, v_1 \rangle = ... = \langle v, v_n \rangle = 0\}.$$

(v) Sejam U_1, U_2 subespaços de V. Tem-se

$$(U_1 + U_2)^{\perp} = (U_1)^{\perp} \cap (U_2)^{\perp}$$

e

$$(U_1 \cap U_2)^{\perp} \supset (U_1)^{\perp} + (U_2)^{\perp}.$$

Se $\dim V < \infty$ tem-se

$$(U_1 \cap U_2)^{\perp} = (U_1)^{\perp} + (U_2)^{\perp}.$$

Dem. (v) Como

$$U_1 \subset U_1 + U_2 \quad \text{e} \quad U_2 \subset U_1 + U_2$$

então

$$(U_1 + U_2)^{\perp} \subset (U_1)^{\perp} \quad \text{e} \quad (U_1 + U_2)^{\perp} \subset (U_2)^{\perp}.$$

Logo

$$(U_1 + U_2)^{\perp} \subset (U_1)^{\perp} \cap (U_2)^{\perp}$$

e assim
$$\left((U_1)^\perp \cap (U_2)^\perp\right)^\perp \subset \left((U_1+U_2)^\perp\right)^\perp$$
e substituindo U_1 por $(U_1)^\perp$ e U_2 por $(U_2)^\perp$ tem-se
$$(U_1 \cap U_2)^\perp \subset \left(\left((U_1)^\perp + (U_2)^\perp\right)^\perp\right)^\perp \underset{\text{Se } \dim V<\infty}{=} (U_1)^\perp + (U_2)^\perp.$$
Por outro lado, como
$$U_1 \cap U_2 \subset U_1 \quad \text{e} \quad U_1 \cap U_2 \subset U_2$$
então
$$(U_1)^\perp \subset (U_1 \cap U_2)^\perp \quad \text{e} \quad (U_2)^\perp \subset (U_1 \cap U_2)^\perp.$$
Logo
$$(U_1)^\perp + (U_2)^\perp \subset (U_1 \cap U_2)^\perp$$
e assim
$$\left((U_1 \cap U_2)^\perp\right)^\perp \subset \left((U_1)^\perp + (U_2)^\perp\right)^\perp$$
e substituindo U_1 por $(U_1)^\perp$ e U_2 por $(U_2)^\perp$ tem-se
$$(U_1)^\perp \cap (U_1)^\perp \subset \left(\left((U_1)^\perp \cap (U_1)^\perp\right)^\perp\right)^\perp \subset (U_1+U_2)^\perp.$$

Exemplo 43. **(i)** Se $U \subset \mathbb{R}^3$ é um plano que passa pela origem, então $U^\perp$ é uma recta que passa pela origem e é perpendicular ao plano.

(ii) Se $U \subset \mathbb{R}^3$ é uma recta que passa pela origem, então $U^\perp$ é um plano que passa pela origem e é perpendicular à recta.

(iii) Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Então, usando o produto interno usual, tem-se
$$(\mathcal{L}(A))^\perp = \mathcal{N}(A).$$

(iv) Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Então, usando um produto interno não necessariamente usual relativamente a uma base $\mathcal{B}$, tem-se
$$(\mathcal{L}(A))^\perp = \mathcal{N}(AG_{\mathcal{B}}).$$

(v) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é invertível. Então, $(\mathcal{N}(A))^\perp = \mathbb{R}^n$ e $(\mathcal{L}(A))^\perp = \{\mathbf{0}\}$.

Definição 55. Sejam U e V espaços lineares. Diz-se que $T : U \to V$ é uma **transformação linear** se e só se verificar as duas condições:

(i) $T(u+v) = T(u) + T(v)$, para todos os $u, v \in U$.

(ii) $T(\lambda u) = \lambda T(u)$, para todos os $u \in U$ e escalares λ.

Definição 56. Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear. Seja $\mathbf{0}$ o vector nulo de V.

(i) Chama-se **contradomínio** ou imagem de T ao conjunto

$$\mathcal{I}(T) = \{T(u) : u \in U\}.$$

Note-se que $\mathcal{I}(T)$ é um subespaço linear de V.

(ii) Chama-se **núcleo** ou espaço nulo de T ao conjunto

$$\mathcal{N}(T) = \{u \in U : T(u) = \mathbf{0}\}.$$

Note-se que $\mathcal{N}(T)$ é um subespaço linear de U.

Observação 32. Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear. Iremos ver que se U tiver dimensão finita então

$$\dim \mathcal{N}(T) + \dim \mathcal{I}(T) = \dim U.$$

Teorema 58. Se U é um subespaço de um espaço euclidiano V, então V é a soma directa de U e $U^\perp$, isto é,

$$V = U \oplus U^\perp.$$

Logo, cada elemento $v \in V$ pode ser escrito de modo único como soma de um elemento de U com um elemento de $U^\perp$:

$$v = v_U + v_{U^\perp}, \quad \text{com} \quad v_U \in U \quad \text{e} \quad v_{U^\perp} \in U^\perp.$$

À transformação linear $P_U : V \to V$ definida por

$$P_U(v) = v_U$$

chama-se **projecção ortogonal de** V **sobre** U. Note que P_U satisfaz

$$P_U = P_U \circ P_U = (P_U)^2 \qquad \text{e} \qquad P_U(v) = \begin{cases} v & \text{se } v \in U \\ \mathbf{0} & \text{se } v \in U^\perp \end{cases}$$

e

$$\mathcal{I}(P_U) = U \qquad \text{e} \qquad \mathcal{N}(P_U) = U^\perp.$$

À transformação linear $P_{U^\perp} : V \to V$ definida por

$$P_{U^\perp}(v) = v_{U^\perp}$$

chama-se **projecção ortogonal de** V **sobre** $U^\perp$. Note que $P_{U^\perp}$ satisfaz

$$P_{U^\perp} = P_{U^\perp} \circ P_{U^\perp} = (P_{U^\perp})^2 \qquad \text{e} \qquad P_{U^\perp}(v) = \begin{cases} v & \text{se } v \in U^\perp \\ \mathbf{0} & \text{se } v \in U \end{cases}$$

e

$$\mathcal{I}\left(P_{U^\perp}\right) = U^\perp \qquad \text{e} \qquad \mathcal{N}\left(P_{U^\perp}\right) = U.$$

Tem-se

$$I = P_U + P_{U^\perp} \qquad \text{e} \qquad \dim V = \dim U + \dim U^\perp$$

uma vez que, como veremos a seguir, tem-se $\dim V = \dim \mathcal{N}(T) + \dim \mathcal{I}(T)$, para qualquer transformação linear T definida num espaço linear de dimensão finita. Logo

$$\left(U^\perp\right)^\perp = U$$

Se $\{w_1, w_2, ..., w_l\}$ fôr uma base ortogonal de U, então

$$P_U(v) = \sum_{i=1}^{l} \frac{\langle v, w_i \rangle}{\|w_i\|^2} w_i = \sum_{i=1}^{l} \operatorname{proj}_{w_i} v = v_U$$

para todo o $v \in V$.

Se $\{u_1, u_2, ..., u_k\}$ é uma base ortogonal de $U^\perp$, então, para todo o $v \in V$

$$P_{U^\perp}(v) = \sum_{j=1}^{k} \frac{\langle v, u_j \rangle}{\|u_j\|^2} u_j = \sum_{j=1}^{k} \operatorname{proj}_{u_j} v = v_{U^\perp}$$

Neste caso, $\{w_1, w_2, ..., w_l, u_1, u_2, ..., u_k\}$ é uma base ortogonal de V.

Tem-se ainda:

(i)

$$\langle P_U(u), v \rangle = \langle u, P_U(v) \rangle, \qquad \langle P_{U^\perp}(u), v \rangle = \langle u, P_{U^\perp}(v) \rangle,$$

para todos os $u, v \in V$;

(ii)

$$\|u\|^2 = \|P_U(u)\|^2 + \|P_{U^\perp}(u)\|^2,$$

para todo o $u \in V$ (Teorema de Pitágoras);

Teorema 59. Seja U um subespaço de um espaço euclidiano V. Seja $v \in V$. Então, tem-se

$$\|v - P_U(v)\| \leq \|v - u\|,$$

para todo o $u \in U$, e a igualdade verifica-se se e só se $u = P_U(v)$.

Dem.

$$\|v - P_U(v)\|^2 = \|v - u - P_U(v-u)\|^2 = \|P_{U^\perp}(v-u)\|^2 \leq$$
$$\leq \|P_{U^\perp}(v-u)\|^2 + \|P_U(v-u)\|^2 \underset{\text{Pitágoras}}{=}$$

$$= \|v-u\|^2 \Leftrightarrow \|v-P_U(v)\| \leq \|v-u\|.$$

A igualdade $\|v-P_U(v)\| = \|v-u\|$ verifica-se se e só $P_U(v-u)$, isto é, se e só se $u = P_U(v)$.

Definição 57. Seja U um subespaço de dimensão finita de um espaço linear V com produto interno. Seja $v \in V$. Então, **o elemento de U mais próximo de v é a projecção ortogonal $P_U(v)$ de v sobre U.**

Definição 58. Seja U um subespaço de um espaço euclidiano V. A **distância d entre um ponto $v \in V$ e um subespaço U** é dada por:

$$d(v,U) = \|P_{U^\perp}(v-\mathbf{0})\| = \|P_{U^\perp}(v)\| = \|v-P_U(v)\|.$$

Definição 59. Seja V um espaço euclidiano. Seja U um subespaço de V com $\dim U = k$. Seja $q \in V$. Chama-se ao conjunto

$$\{q\}+U$$

um k-plano. A **distância d entre um ponto $p \in V$ e um k-plano $\mathcal{P} = \{q\}+U$** é dada por:

$$d(p,\mathcal{P}) = \|P_{U^\perp}(p-q)\|.$$

Definição 60. **(i)** A distância entre dois k-planos paralelos

$$\mathcal{P}_1 = \{p\}+U \quad \text{e} \quad \mathcal{P}_2 = \{q\}+U$$

é dada por:

$$d(\mathcal{P}_1,\mathcal{P}_2) = \|P_{U^\perp}(p-q)\|.$$

(ii) A distância entre duas rectas paralelas

$$r = \{p\}+L(\{u\}) \quad \text{e} \quad s = \{q\}+L(\{u\})$$

é dada por:

$$d(r,s) = \left\| P_{L(\{u\})^\perp}(p-q) \right\|.$$

Exemplo 44. Considere-se $\mathbb{R}^3$ com o produto interno usual.

(i) Seja $\mathcal{P}$ o plano (em $\mathbb{R}^3$) que passa pelos pontos: $(1,2,1)$, $(1,0,-1)$ e $(1,1,1)$. Tem-se

$$\mathcal{P} = \{(1,2,1)\} + L(\{(0,-2,-2),(0,-1,0)\})$$

uma vez que

$$(0,-2,-2) = (1,0,-1)-(1,2,1) \qquad \text{e} \qquad (0,-1,0) = (1,1,1)-(1,2,1).$$

Equação vectorial de $\mathcal{P}$: $(x, y, z) = (1, 2, 1) + \alpha(0, -2, -2) + \beta(0, -1, 0)$, com $\alpha, \beta \in \mathbb{R}$.

Equações paramétricas de $\mathcal{P}$:

$$\begin{cases} x = 1 \\ y = 2 - \beta - 2\alpha \\ z = 1 - 2\alpha \end{cases}$$

com $\alpha, \beta \in \mathbb{R}$.

Equação cartesiana de $\mathcal{P}$: $x = 1$.

Podemos determinar a **equação cartesiana de $\mathcal{P}$** do seguinte modo. Atendendo a que

$$\mathcal{P} = \{(1, 2, 1)\} + L\left(\{(0, -2, -2), (0, -1, 0)\}\right)$$

seja

$$U = L\left(\{(0, -2, -2), (0, -1, 0)\}\right).$$

Logo,

$$U = \left(U^{\perp}\right)^{\perp} = \left(\mathcal{N}\left(\begin{bmatrix} 0 & -2 & -2 \\ 0 & -1 & 0 \end{bmatrix}\right)\right)^{\perp} =$$

$$= \left(L\left(\{(1, 0, 0)\}\right)\right)^{\perp} = \left\{(x, y, z) \in \mathbb{R}^3 : \langle (x, y, z), (1, 0, 0) \rangle = 0\right\}$$

e assim, a equação cartesiana do plano $\mathcal{P}$ que passa pelo ponto $(1, 2, 1)$ é dada por:

$$\left(\langle (x - 1, y - 2, z - 1), (1, 0, 0) \rangle = 0\right) \Leftrightarrow \left(1\,(x - 1) + 0\,(y - 2) + 0\,(z - 1) = 0\right),$$

ou seja por

$$x = 1.$$

(ii) Determinemos a **equação cartesiana** da recta que passa pelos pontos $(1, 1, 0)$ e $(1, 2, 1)$. Tem-se

$$r = \{(1, 1, 0)\} + L\left(\{(0, 1, 1)\}\right),$$

uma vez que

$$(0, 1, 1) = (1, 2, 1) - (1, 1, 0).$$

Seja

$$U = L\left(\{(0, 1, 1)\}\right).$$

Logo,

$$U = \left(U^{\perp}\right)^{\perp} = \left(\mathcal{N}\left(\begin{bmatrix} 0 & 1 & 1 \end{bmatrix}\right)\right)^{\perp} = \left(L\left(\{(1, 0, 0), (0, 1, -1)\}\right)\right)^{\perp}$$

e assim, a equação cartesiana da recta r é dada por:

$$\left(\langle (x - 1, y - 1, z), (1, 0, 0) \rangle = 0 \text{ e } \langle (x - 1, y - 1, z), (0, 1, -1) \rangle = 0\right) \Leftrightarrow$$

$$\Leftrightarrow \left(1\,(x - 1) = 0 \text{ e } 1\,(y - 1) - 1z = 0\right),$$

ou seja por

$$\begin{cases} x = 1 \\ y - z = 1. \end{cases}$$

Transformações lineares

Definição 61. Sejam U e V espaços lineares. Diz-se que

$$T : U \to V$$

é uma **transformação linear** se e só se verificar as duas condições:

(i) $T(u + v) = T(u) + T(v)$, para todos os $u, v \in U$.

(ii) $T(\lambda u) = \lambda T(u)$, para todos os $u \in U$ e escalares λ.

Observação 33. Sejam U e V espaços lineares. Sejam $\mathbf{0}$ o vector nulo de U e $\mathbf{0}'$ o vector nulo de V.

(i) Se $T : U \to V$ fôr uma transformação linear então $T(U)$ é um subespaço de V e além disso tem-se $T(\mathbf{0}) = \mathbf{0}'$ ($T(\mathbf{0}) = T(\mathbf{0} + \mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0}) \Leftrightarrow T(\mathbf{0}) = \mathbf{0}'$). Logo, se T não verificar $T(\mathbf{0}) = \mathbf{0}'$ então T não será uma transformação linear.

(ii) $T : U \to V$ é uma transformação linear se e só se

$$T(\lambda u + \mu v) = \lambda T(u) + \mu T(v),$$

para todos os $u, v \in U$ e escalares λ, μ.

(iii) Seja $T : U \to V$ uma transformação linear, com $U = L\left(\{v_1, ..., v_n\}\right)$. Seja $u \in U$. Logo, existem escalares $\lambda_1, ..., \lambda_n$ tais que

$$u = \lambda_1 v_1 + ... + \lambda_n v_n.$$

Tem-se então

$$T(u) = \lambda_1 T(v_1) + ... + \lambda_n T(v_n).$$

Exemplo 45. Consideremos a base canónica $\{(1, 0), (0, 1)\}$ de $\mathbb{R}^2$. Seja

$$T : \mathbb{R}^2 \to \mathbb{R}$$

uma transformação linear tal que $T(1, 0) = 1$ e $T(0, 1) = 1$.

Para qualquer $(x, y) \in \mathbb{R}^2$ tem-se

$$(x, y) = x(1, 0) + y(0, 1).$$

Então,

$$T(x, y) = T\left(x(1, 0) + y(0, 1)\right) = xT(1, 0) + yT(0, 1) = x + y.$$

Logo, $T : \mathbb{R}^2 \to \mathbb{R}$ é a transformação linear definida explicitamente por

$$T(x, y) = x + y.$$

Teorema 60. Sejam U e V espaços lineares e seja $\{v_1, ..., v_n\}$ uma base de U. Sejam $T_1, T_2 : U \to V$ duas transformações lineares.

$$\text{Se} \quad T_1(v_i) = T_2(v_i) \quad \text{para todo o } i = 1, ..., n, \text{ então} \quad T_1(u) = T_2(u),$$

para todo o $u \in U$, isto é, $T_1 = T_2$.

Exemplo 46. **(i)** Sejam U e V espaços lineares e seja $\mathbf{0}$ o vector nulo de V. Seja $O : U \to V$ definida por

$$O(u) = \mathbf{0},$$

para todo o $u \in U$. O é uma transformação linear e chama-se **transformação nula**.

(ii) Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Seja

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

definida por

$$T(u) = Au,$$

para todo o $u \in \mathbb{R}^n$. T é uma transformação linear.

(iii) Sejam V um espaço linear e k um escalar (fixo). Seja $T_k : V \to V$ definida por

$$T_k(v) = kv,$$

para todo o $v \in V$.

T_k é uma transformação linear. Diz-se que T_k é uma **homotetia**.

Se $0 < k < 1$ diz-se que T_k é uma **contracção**.

Se $k > 1$ diz-se que T_k é uma **dilatação**.

Se $k = 1$ então chama-se a T_1 a **transformação identidade** e denota-se por I. Tem-se

$$I(u) = u,$$

para todo o $u \in U$.

(iv) $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (1 - y, 2x)$ **não** é uma transformação linear.

(v) $T : \mathbb{R}^2 \to \mathbb{R}$ definida por $T(x, y) = xy$ **não** é uma transformação linear.

(vi) Seja $T : \mathcal{P}_2 \to \mathcal{P}_3$ definida por

$$T(p(t)) = tp(t).$$

T é uma transformação linear.

(vii) Seja $T : \mathcal{P}_3 \to \mathcal{P}_1$ definida por

$$T(p) = p''.$$

T é uma transformação linear.

(viii) Seja $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ definida por

$$T(f) = f',$$

onde $C^1(\mathbb{R})$ é o espaço linear de todas as funções reais com primeira derivada contínua em $\mathbb{R}$ e $C(\mathbb{R})$ é o espaço linear de todas as funções reais contínuas em $\mathbb{R}$. T é uma transformação linear.

(ix) Seja $a \in \mathbb{R}$ (fixo). Seja $T : C^1(\mathbb{R}) \to \mathbb{R}$ definida por

$$T(f) = f'(a).$$

T é uma transformação linear.

(x) Seja $n \in \mathbb{N}$. Seja $T : C^n(\mathbb{R}) \to C(\mathbb{R})$ definida por

$$T(f) = f^{(n)},$$

onde $f^{(n)}$ é a derivada de ordem n de f, $C^n(\mathbb{R})$ é o espaço linear de todas as funções reais com derivada de ordem n contínua em $\mathbb{R}$ e $C(\mathbb{R})$ é o espaço linear de todas as funções reais contínuas em $\mathbb{R}$. T é uma transformação linear.

(xi) Seja $T : C(\mathbb{R}) \to C^1(\mathbb{R})$ definida por

$$T(f) = \int_0^x f(t)\,dt.$$

T é uma transformação linear.

(xii) Seja $T : C([a, b]) \to \mathbb{R}$ definida por

$$T(f) = \int_a^b f(t)\,dt.$$

T é uma transformação linear.

(xiii) Seja $T : \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathcal{M}_{n\times n}(\mathbb{R})$ definida por

$$T(X) = X^T.$$

T é uma transformação linear.

(xiv) Seja $T : \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathcal{M}_{n\times n}(\mathbb{R})$ definida por

$$T(X) = AX,$$

com $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ fixa. T é uma transformação linear.

(xv) Seja

$$\mathrm{tr} : \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathbb{R}$$

definida por

$$\mathrm{tr}(A) = a_{11} + ... + a_{nn} = \sum_{i=1}^{n} a_{ii},$$

para todo o $A = (a_{ij})_{n\times n} \in \mathcal{M}_{n\times n}(\mathbb{R})$, isto é, $\operatorname{tr}(A)$ é a soma de todas as entradas da diagonal principal de A. O **traço**, tr, é uma transformação linear, isto é, sendo $A = (a_{ij})_{n\times n}$ e $B = (b_{ij})_{n\times n}$ duas matrizes do tipo $n \times n$ e α um escalar, tem-se

$$\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B) \qquad \text{e} \qquad \operatorname{tr}(\alpha A) = \alpha \operatorname{tr}(A).$$

Além disso, tem-se

$$\operatorname{tr}(A^T) = \operatorname{tr}(A) \qquad \text{e} \qquad \operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Definição 62. Sejam U e V espaços lineares e $T_1, T_2 : U \to V$ transformações lineares. Seja λ um escalar. Sejam $T_1 + T_2, \lambda T_1 : U \to V$ definidas por

$$(T_1 + T_2)(u) = T_1(u) + T_2(u) \qquad \text{e} \qquad (\lambda T_1)(u) = \lambda T_1(u),$$

para todo o $u \in U$.

Definição 63. Sejam U e V espaços lineares. Chama-se a $\mathfrak{L}(U, V)$ o conjunto de **todas** as transformações lineares de U em V.

Teorema 61. Sejam U e V espaços lineares e $T_1, T_2 : U \to V$ transformações lineares. Seja λ um escalar. Então:

(i) $T_1 + T_2$ e λT_1 são transformações lineares.

(ii) O conjunto $\mathfrak{L}(U, V)$, com as operações da definição 38, é um espaço linear.

Exemplo 47. Seja $\mathcal{B} = \{T_1, T_2, T_3, T_4\}$ com $T_1, T_2, T_3, T_4 \in \mathfrak{L}(\mathbb{R}^2, \mathbb{R}^2)$ definidas por

$$T_1(x, y) = (x, 0), \;\; T_2(x, y) = (y, 0), \;\; T_3(x, y) = (0, x) \;\; \text{e} \;\; T_4(x, y) = (0, y),$$

para todo o $(x, y) \in \mathbb{R}^2$. O conjunto $\mathcal{B}$ é uma base de $\mathfrak{L}(\mathbb{R}^2, \mathbb{R}^2)$. Logo,

$$\dim \mathfrak{L}(\mathbb{R}^2, \mathbb{R}^2) = 4.$$

Definição 64. Sejam U, V e W espaços lineares e, $T_2 : U \to V$ e $T_1 : V \to W$ transformações lineares. Seja $T_1 \circ T_2 : U \to W$ definida por

$$(T_1 \circ T_2)(u) = T_1(T_2(u)),$$

para todo o $u \in U$. Chama-se a $T_1 \circ T_2$ a **composição de** T_1 com T_2.

Observação 34. Em geral, tem-se $T_1 \circ T_2 \neq T_2 \circ T_1$.

Teorema 62. **(i)** Sejam $T_2 : U \to V$ e $T_1 : V \to W$ transformações lineares. Então $T_1 \circ T_2$ é uma transformação linear.

(ii) Sejam $T_3 : U \to V$, $T_2 : V \to W$ e $T_1 : W \to X$. Então, tem-se $T_1 \circ (T_2 \circ T_3) = (T_1 \circ T_2) \circ T_3$.

(iii) Sejam $T_4 : W \to U$, $T_2, T_3 : U \to V$ e $T_1 : V \to W$. Seja $\lambda \in \mathbb{R}$. Então, tem-se

$$T_1 \circ (T_2 + T_3) = T_1 \circ T_2 + T_1 \circ T_3 \;\text{ e }\; T_1 \circ (\lambda T_2) = \lambda (T_1 \circ T_2).$$

$$(T_2 + T_3) \circ T_4 = T_2 \circ T_4 + T_3 \circ T_4 \;\text{ e }\; (\lambda T_3) \circ T_4 = \lambda (T_3 \circ T_4).$$

Definição 65 Define-se $T^0 = I$ e $T^k = T \circ T^{k-1}$, para todo o $k = 1, 2, ...$

Observação 35. Tem-se $T^{m+n} = T^m \circ T^n$ para todos os $m, n \in \mathbb{N}$.

Definição 66. Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear. Seja $\mathbf{0}$ o vector nulo de V.

(i) Chama-se **contradomínio** ou imagem de T ao conjunto

$$T(U) = \{T(u) : u \in U\},$$

que também se denota por $\mathcal{I}(T)$.

Note-se que se existir $\{u_1, ..., u_k\} \subset U$ tal que $U = L(\{u_1, ..., u_k\})$ então

$$\mathcal{I}(T) = L(\{T(u_1), ..., T(u_k)\}).$$

(ii) Chama-se **núcleo** ou espaço nulo de T ao conjunto

$$\mathcal{N}(T) = \{u \in U : T(u) = \mathbf{0}\}.$$

Teorema 63. Sejam U e V espaços lineares e $T : U \to V$ uma transformação linear. Então, os conjuntos $\mathcal{N}(T)$ e $\mathcal{I}(T)$ são subespaços de U e V respectivamente.

Exemplo 48. **(i)** Sejam U e V espaços lineares. Sejam $\mathbf{0}$ e $\mathbf{0}'$ os vectores nulos de U e V respectivamente.

Considere a transformação nula $O : U \to V$ definida por

$$O(u) = \mathbf{0}',$$

para todo o $u \in U$. Tem-se

$$\mathcal{N}(O) = U \;\text{ e }\; \mathcal{I}(O) = \{\mathbf{0}'\}.$$

(ii) Considere a transformação identidade $I : U \to U$ definida por

$$I(u) = u,$$

para todo o $u \in U$. Tem-se

$$\mathcal{N}(I) = \{\mathbf{0}\} \quad \text{e} \quad \mathcal{I}(I) = U.$$

(iii) Seja $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. Seja

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

definida por

$$T(u) = Au,$$

para todo o $u \in \mathbb{R}^n$. Tem-se

$$\mathcal{N}(T) = \mathcal{N}(A) \quad \text{e} \quad \mathcal{I}(T) = \mathcal{C}(A).$$

(iv) Seja $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ definida por

$$T(f) = f'.$$

Tem-se

$$\mathcal{N}(T) = \{f : \mathbb{R} \to \mathbb{R} \text{ tal que } f \text{ é constante em } \mathbb{R}\} \qquad \text{e} \qquad \mathcal{I}(T) = C(\mathbb{R}).$$

(v) Seja $T : C^2(\mathbb{R}) \to C(\mathbb{R})$ definida por

$$T(f(t)) = f''(t) + \omega^2 f(t),$$

com $\omega \in \mathbb{R} \backslash \{0\}$. Tem-se (pág. 72 de [1])

$$\mathcal{N}(T) = L(\{\cos(\omega t), \operatorname{sen}(\omega t)\}),$$

onde $\{\cos(\omega t), \operatorname{sen}(\omega t)\}$ é uma base de $\mathcal{N}(T)$. Observe-se que $\mathcal{N}(T)$ é precisamente a solução geral da equação diferencial linear homogénea

$$f''(t) + \omega^2 f(t) = 0.$$

(vi) Seja $T : C^2(\mathbb{R}) \to C(\mathbb{R})$ definida por

$$T(f(t)) = f''(t) - \omega^2 f(t),$$

com $\omega \in \mathbb{R} \backslash \{0\}$. Tem-se (pág. 74 de [1])

$$\mathcal{N}(T) = L\left(\left\{e^{-\omega t}, e^{\omega t}\right\}\right),$$

onde $\{e^{-\omega t}, e^{\omega t}\}$ é uma base de $\mathcal{N}(T)$. Note-se que $\mathcal{N}(T)$ é precisamente a solução geral da equação diferencial linear homogénea

$$f''(t) - \omega^2 f(t) = 0.$$

Definição 67. $T : U \to V$ diz-se **injectiva** se e só se

$$T(u) = T(w) \quad \Rightarrow \quad u = w,$$

para todos os $u, w \in U$, isto é, se e só se

$$u \neq w \quad \Rightarrow \quad T(u) \neq T(w),$$

para todos os $u, w \in U$.

Teorema 64. **(i)** Sejam U e V espaços lineares. Seja $T : U \to V$ uma qualquer transformação linear. Então:

$$T \text{ é injectiva } \Leftrightarrow \mathcal{N}(T) = \{\mathbf{0}\}.$$

(ii) Sejam U um espaço linear de dimensão finita e T uma transformação linear definida em U. Então, o subespaço $\mathcal{I}(T)$ tem dimensão finita e

$$\dim \mathcal{N}(T) + \dim \mathcal{I}(T) = \dim U.$$

Dem. **(i)** **($\Rightarrow$)** Suponhamos que T é injectiva. Seja $u \in \mathcal{N}(T)$. Logo $T(u) = \mathbf{0}_V \underset{T \text{ é linear}}{=} T(\mathbf{0}_U)$, pelo que $u = \mathbf{0}$ uma vez que T é injectiva. Logo $\mathcal{N}(T) = \{\mathbf{0}\}$.

($\Leftarrow$) Suponhamos que $\mathcal{N}(T) = \{\mathbf{0}\}$. Sejam $u, v \in U$ tais que $T(u) = T(v)$. Logo $T(u - v) = \mathbf{0}$, pelo que $u - v = \mathbf{0}$ uma vez que $\mathcal{N}(T) = \{\mathbf{0}\}$. Logo $u = v$ e assim T é injectiva.

(ii) Se $\dim \mathcal{N}(T) = 0$ então T é injectiva, pela alínea (i). Suponhamos que $\dim U = n$. Considerando uma base $\{w_1, ..., w_n\}$ de U, vamos mostrar que o conjunto de n vectores $\{T(w_1), ..., T(w_n)\}$ é uma base de $\mathcal{I}(T)$.

Seja $v \in \mathcal{I}(T)$. Existe então $u \in U$ tal que $v = T(u)$. Como $\{w_1, ..., w_n\}$ é base de U, existem escalares (únicos) $\alpha_1, ..., \alpha_n$ tais que $u = \sum_{i=1}^{n} \alpha_i w_i$. Logo, como T é linear, $v = T(u) = \sum_{i=1}^{n} \alpha_i T(w_i)$ concluindo-se deste modo que o conjunto $\{T(w_1), ..., T(w_n)\}$ gera $\mathcal{I}(T)$.

Sejam agora $\beta_1, ..., \beta_n$ escalares tais que $\beta_1 T(w_1) + ... + \beta_n T(w_n) = \mathbf{0}$. A última igualdade é equivalente a $T(\beta_1 w_1 + ... + \beta_n w_n) = \mathbf{0}$ uma vez que T é linear. Logo, como T é injectiva, obtém-se $\beta_1 w_1 + ... + \beta_n w_n = \mathbf{0}$ e, deste modo, $\beta_1 = ... = \beta_n = 0$ uma vez que o conjunto $\{w_1, ..., w_n\}$ é linearmente independente.

Seja $n = \dim U$. Suponhamos agora que $\dim \mathcal{N}(T) \neq 0$. Seja $r = \dim \mathcal{N}(T)$ e seja $\{u_1, ..., u_r\}$ uma base de $\mathcal{N}(T)$. Considere-se os vectores $u_{r+1}, ..., u_n \in U$ de modo a que $\{u_1, ..., u_r, u_{r+1}, ..., u_n\}$ seja uma base de U. Vejamos que $\{T(u_{r+1}), ..., T(u_n)\}$ é uma base de $\mathcal{I}(T)$.

Seja $v \in \mathcal{I}(T)$. Existe então $u \in U$ tal que $v = T(u)$. Como $\{u_1, ..., u_n\}$ é base de U, existem escalares (únicos) $\alpha_1, ..., \alpha_n$ tais que $u = \sum_{i=1}^{n} \alpha_i u_i$. Logo, como T é linear, $v = T(u) = \sum_{i=1}^{n} \alpha_i T(u_i)$ concluindo-se deste modo que o conjunto

$$\{T(u_1), ..., T(u_n)\} \underset{\{u_1,...,u_r\} \in \mathcal{N}(T)}{=} \{T(u_{r+1}), ..., T(u_n)\}$$

gera $\mathcal{I}(T)$.

Sejam agora $\beta_{r+1}, ..., \beta_n$ escalares tais que $\beta_{r+1} T(u_{r+1}) + ... + \beta_n T(u_n) = \mathbf{0}$. A última igualdade é equivalente a $T\left(\beta_{r+1} u_{r+1} + ... + \beta_n u_n\right) = \mathbf{0}$ uma vez que T é linear. Logo $\beta_{r+1} u_{r+1} + ... + \beta_n u_n \in \mathcal{N}(T)$. Por outro lado, como $\{u_1, ..., u_r\}$ é base de $\mathcal{N}(T)$, existem escalares (únicos) $\gamma_1, ..., \gamma_r$ tais que

$$\beta_{r+1} u_{r+1} + ... + \beta_n u_n = \sum_{i=1}^{r} \gamma_i u_i,$$

ou seja

$$\sum_{i=1}^{r} (-\gamma_i) u_i + \sum_{i=r+1}^{n} \beta_i u_i = \mathbf{0}$$

de onde se obtem

$$\gamma_1 = ... = \gamma_r = \beta_{r+1} = ... = \beta_n = 0$$

uma vez que o conjunto $\{u_1, ..., u_r\}$ é linearmente independente. Assim $\beta_{r+1} = ... = \beta_n = 0$ e deste modo, o conjunto $\{T(u_{r+1}), ..., T(u_n)\}$ é linearmente independente.

Exemplo 49. Seja $A \in \mathcal{M}_{m \times n}(\mathbb{R})$. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ tal que $T(u) = Au$, para todo o $u \in \mathbb{R}^n$. Tem-se

$$\mathcal{N}(A) = \mathcal{N}(T) \quad \text{e} \quad \mathcal{C}(A) = \mathcal{I}(T)$$

e também:

$$\operatorname{car} A + \operatorname{nul} A = \dim \mathcal{C}(A) + \dim \mathcal{N}(A) = \dim \mathcal{I}(T) + \dim \mathcal{N}(t) = \dim \mathbb{R}^n = n.$$

Definição 68. **(i)** $T : U \to V$ diz-se **sobrejectiva** se e só se $T(U) = V$.

(ii) $T : U \to V$ diz-se **bijectiva** se e só se fôr injectiva e sobrejectiva.

Definição 69. Sejam U e V espaços lineares. Diz-se que U e V são isomorfos se e só se existir um **isomorfismo** entre U e V, isto é, se e só se existir uma transformação linear bijectiva $T : U \to V$. Sendo U e V isomorfos escreve-se

$$U \cong V.$$

Teorema 65. Sejam U e V dois espaços lineares de dimensões finitas. Então, U e V são isomorfos se e só se $\dim U = \dim V$.

Teorema 66. **(i)** Qualquer espaço linear real de dimensão n é isomorfo a $\mathbb{R}^n$.

(ii) Sejam U e V dois espaços lineares de dimensões finitas. A transformação linear $T : U \to V$ é sobrejectiva se e só se T transformar um qualquer conjunto gerador de U num conjunto gerador de V.

(iii) Sejam U e V dois espaços lineares de dimensões finitas. Se a transformação linear $T : U \to V$ fôr sobrejectiva então $\dim V \leq \dim U$.

(iv) Sejam U e V dois espaços lineares de dimensões finitas. Se a transformação linear $T : U \to V$ fôr injectiva então $\dim U \leq \dim V$.

Exemplo 50. **(i)** A transformação linear $T : \mathbb{R}^n \to \mathcal{M}_{n\times 1}(\mathbb{R})$ definida por

$$T(a_1, ..., a_n) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix},$$

é um isomorfismo. Logo

$$\mathbb{R}^n \cong \mathcal{M}_{n\times 1}(\mathbb{R}).$$

(ii) A transformação linear $T : \mathcal{M}_{m\times n}(\mathbb{R}) \to \mathbb{R}^{mn}$ definida por

$$T\left(\begin{bmatrix} a_{11} & & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & & a_{mn} \end{bmatrix}\right) = (a_{11}, ..., a_{m1}, ..., a_{1n}, ..., a_{mn}),$$

é um isomorfismo. Logo

$$\mathcal{M}_{m\times n}(\mathbb{R}) \cong \mathbb{R}^{mn}.$$

(iii) A transformação linear $T : \mathbb{R}^{n+1} \to \mathcal{P}_n$ definida por

$$T(a_0, a_1, ..., a_n) = a_0 + a_1 t + ... + a_n t^n,$$

é um isomorfismo. Logo $\mathbb{R}^{n+1} \cong \mathcal{P}_n$.

(iv) Seja A uma matriz $m \times n$. Os espaços $\mathcal{C}(A)$ e $\mathcal{L}(A)$ são isomorfos pois têm a mesma dimensão ($\operatorname{car} A$).

$$\mathcal{C}(A) \cong \mathcal{L}(A).$$

Observação 36. Sejam U e V espaços lineares de dimensões finitas tais que

$$\dim U = \dim V.$$

Seja $T : U \to V$ uma transformação linear. Então, T é injectiva se e só se T é sobrejectiva.

Definição 70. Sejam U e V espaços lineares de dimensões finitas tais que $\dim U = \dim V$. Diz-se que $T : U \to V$ é invertível se existir $S : V \to U$ tal que

$$S \circ T = I_U \quad \text{e} \quad T \circ S = I_V,$$

onde I_U é a transformação identidade em U e I_V é a transformação identidade em V. Chama-se a S a inversa de T e escreve-se

$$S = T^{-1}.$$

Se T fôr linear então, a existir, T^{-1} também é linear.

Teorema 67. Sejam U e V espaços lineares de dimensões finitas. Seja

$$T : U \to V$$

uma transformação linear. Seja $\mathbf{0}$ o vector nulo de U. As seguintes afirmações são equivalentes.

(i) T é injectiva.

(ii) $\mathcal{N}(T) = \{\mathbf{0}\}$.

(iii) $\dim U = \dim T(U)$.

(iv) T transforma vectores linearmente independentes de U em vectores linearmente independentes de V.

(v) T transforma bases de U em bases de $T(U)$.

Teorema 68. Sejam U e V espaços lineares. Seja $T : U \to V$ uma transformação linear. Seja $b \in V$. Então:

(i) Existência de solução: a equação linear $T(u) = b$ tem sempre solução (para qualquer b) se e só se T fôr sobrejectiva ($T(U) = V$);

(ii) Unicidade de solução: a equação linear $T(u) = b$ a ter solução, ela é única se e só se T fôr injectiva;

(iii) Existência e unicidade de solução: a equação linear $T(u) = b$ tem solução única u se e só se T fôr bijectiva.

Teorema 69. Sejam U e V espaços lineares. Seja $T : U \to V$ uma transformação linear. Seja $b \in V$. A solução geral da equação linear

$$T(u) = b$$

obtém-se somando a uma solução particular dessa equação, a solução geral da equação linear homogénea $T(u) = \mathbf{0}$ ($\mathcal{N}(T)$).

Matriz de mudança de base

Definição 71. Seja V um espaço linear de dimensão n. Sejam $\mathcal{B}_1 = \{v_1, ..., v_n\}$ e $\mathcal{B}_2 = \{w_1, ..., w_n\}$ duas bases ordenadas de V. Seja $S_{\mathcal{B}_1 \to \mathcal{B}_2}$ a matriz cujas colunas são as coordenadas dos vectores de $\mathcal{B}_1$ em relação à base $\mathcal{B}_2$. Isto é,

$$S_{\mathcal{B}_1 \to \mathcal{B}_2} = (s_{ij})_{n \times n} \quad \text{com} \quad v_j = \sum_{i=1}^{n} s_{ij} w_i \quad \text{para todo o } j = 1, ..., n.$$

A matriz $S_{\mathcal{B}_1 \to \mathcal{B}_2}$ é invertível e chama-se **matriz de mudança de base** (da base $\mathcal{B}_1$ para $\mathcal{B}_2$). Tem-se

$$\begin{bmatrix} v_1 & ... & v_n \end{bmatrix} = \begin{bmatrix} w_1 & ... & w_n \end{bmatrix} S_{\mathcal{B}_1 \to \mathcal{B}_2}$$

Teorema 70. Seja V um espaço linear de dimensão n. Sejam $\mathcal{B}_1 = \{v_1, ..., v_n\}$ e $\mathcal{B}_2 = \{w_1, ..., w_n\}$ duas bases ordenadas de V. Seja $S_{\mathcal{B}_1 \to \mathcal{B}_2}$ a matriz cujas colunas são as coordenadas dos vectores de $\mathcal{B}_1$ em relação à base $\mathcal{B}_2$. Isto é,

$$S_{\mathcal{B}_1 \to \mathcal{B}_2} = (s_{ij})_{n \times n} \quad \text{com} \quad v_j = \sum_{i=1}^{n} s_{ij} w_i \quad \text{para todo o } j = 1, ..., n.$$

Se tivermos

$$u = \sum_{i=1}^{n} \lambda_i v_i,$$

isto é, se $(\lambda_1, ..., \lambda_n)$ forem as coordenadas do vector u na base $\mathcal{B}_1$ então as coordenadas $(\mu_1, ..., \mu_n)$ de u na base $\mathcal{B}_2$ são dadas por

$$\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = S_{\mathcal{B}_1 \to \mathcal{B}_2} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

Dem. Tem-se

$$u = \sum_{i=1}^{n} \mu_i w_i = \sum_{j=1}^{n} \lambda_j v_j = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{n} s_{ij} w_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} s_{ij} \lambda_j \right) w_i.$$

como as coordenadas de um vector u numa base são únicas, tem-se para todo o $i = 1, ..., n$,

$$\mu_i = \left(\sum_{j=1}^{n} s_{ij} \lambda_j \right). \quad \text{Isto é,} \quad \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = S_{\mathcal{B}_1 \to \mathcal{B}_2} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

Teorema 71. Tem-se

$$S_{\mathcal{B}_2 \to \mathcal{B}_1} = \left(S_{\mathcal{B}_1 \to \mathcal{B}_2} \right)^{-1}.$$

Exemplo 51. Consideremos

$$\mathcal{B}_c = \{(1,0),(0,1)\}$$

a base canónica de $\mathbb{R}^2$. Seja

$$\mathcal{B} = \{(1,2),(3,4)\}$$

uma outra base ordenada de $\mathbb{R}^2$. Sejam $(5,6)$ as coordenadas de um vector u na base canónica $\mathcal{B}_c$ e determinemos as coordenadas de u na base $\mathcal{B}$ usando a matriz de mudança de base $S_{\mathcal{B}_c \to \mathcal{B}}$. Tem-se

$$S_{\mathcal{B}_c \to \mathcal{B}} = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix},$$

uma vez que

$$(1,0) = -2(1,2) + 1(3,4) \quad \text{e} \quad (0,1) = \frac{3}{2}(1,2) - \frac{1}{2}(3,4). \quad (*)$$

Logo, as coordenadas de u na base $\mathcal{B}$ são dadas por

$$S_{\mathcal{B}_c \to \mathcal{B}} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Logo, -1 e 2 são as coordenadas de $(5,6)$ na base ordenada $\mathcal{B}$, isto é

$$(5,6) = -1(1,2) + 2(3,4).$$

Observação 37. Colocando os vectores em coluna, note que as duas igualdades em (*) podem ser escritas na forma:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}$$

sendo esta última igualdade equivalente a $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}}_{=\begin{bmatrix} -2 & \frac{3}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}^{-1}} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \Leftrightarrow$

$\Leftrightarrow \begin{cases} (1,2) = 1(1,0) + 2(0,1) \\ \\ (3,4) = 3(1,0) + 4(0,1) \end{cases}$ querendo isto dizer que as coordenadas dos vectores $(1,2)$ e $(3,4)$ relativamente à base canónica (ordenada) $\{(1,0),(0,1)\}$ são respectivamente $(1,2)$ e $(3,4)$.

Observação 38. Sendo $\mathcal{B}_1$ e $\mathcal{B}_2$ bases ordenadas de um espaço euclidiano V. Tem-se:

$$\langle u, v \rangle = \left([u]_{\mathcal{B}_1}\right)^T G_{\mathcal{B}_1} [v]_{\mathcal{B}_1}$$

e também

$$\langle u, v \rangle = \left([u]_{\mathcal{B}_2}\right)^T G_{\mathcal{B}_2} [v]_{\mathcal{B}_2}$$

ou seja:

$$\langle u, v \rangle = \left([u]_{\mathcal{B}_2}\right)^T G_{\mathcal{B}_2} [v]_{\mathcal{B}_2} = \left(S_{\mathcal{B}_1 \to \mathcal{B}_2} [u]_{\mathcal{B}_1}\right)^T G_{\mathcal{B}_2} S_{\mathcal{B}_1 \to \mathcal{B}_2} [v]_{\mathcal{B}_1} = \left([u]_{\mathcal{B}_1}\right)^T \left(S^T_{\mathcal{B}_1 \to \mathcal{B}_2} G_{\mathcal{B}_2} S_{\mathcal{B}_1 \to \mathcal{B}_2}\right) [v]_{\mathcal{B}_1}$$

com

$$S^T_{\mathcal{B}_1 \to \mathcal{B}_2} G_{\mathcal{B}_2} S_{\mathcal{B}_1 \to \mathcal{B}_2} = G_{\mathcal{B}_1}.$$

Representação matricial de uma transformação linear

Definição 72. (Representação matricial de uma transformação linear). Sejam U e V espaços lineares de dimensões finitas tais que $\dim U = n$ e $\dim V = m$. Sejam $\mathcal{B}_1 = \{u_1, ..., u_n\}$ e $\mathcal{B}_2 = \{v_1, ..., v_m\}$ duas bases ordenadas de U e V respectivamente. Seja $T : U \to V$ uma transformação linear. Considere-se a matriz $A = (a_{ij})_{m\times n}$ cuja coluna j, para cada $j = 1, ..., n$, é formada pelas coordenadas de $T(u_j)$ na base $\mathcal{B}_2$. Isto é,

$$T(u_j) = \sum_{i=1}^{m} a_{ij} v_i.$$

Chama-se a esta matriz A a **representação matricial** de T em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$ e escreve-se

$$A = M(T; \mathcal{B}_1; \mathcal{B}_2).$$

Teorema 72. (Representação matricial de uma transformação linear). Sejam U e V espaços lineares de dimensões finitas tais que $\dim U = n$ e $\dim V = m$. Sejam $\mathcal{B}_1 = \{u_1, ..., u_n\}$ e $\mathcal{B}_2 = \{v_1, ..., v_m\}$ duas bases ordenadas de U e V respectivamente. Seja $T : U \to V$ uma transformação linear. Seja

$$A = M(T; \mathcal{B}_1; \mathcal{B}_2).$$

Se $\alpha_1, ..., \alpha_n$ forem as coordenadas de um vector $u \in U$ na base ordenada $\mathcal{B}_1$ então as coordenadas $\beta_1, ..., \beta_m$ de $T(u) \in V$ na base ordenada $\mathcal{B}_2$ são dadas por

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} = M(T; \mathcal{B}_1; \mathcal{B}_2) \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \text{isto é:} \qquad [T(u)]_{\mathcal{B}_2} = M(T; \mathcal{B}_1; \mathcal{B}_2) [u]_{\mathcal{B}_1}.$$

Observação 39. MUITO IMPORTANTE. Nas condições do teorema anterior, tem-se

$$u = \sum_{j=1}^{n} \alpha_j u_j \in \mathcal{N}(T) \Leftrightarrow (\alpha_1, ..., \alpha_n) \in \mathcal{N}(A)$$

$$v = \sum_{i=1}^{m} \beta_i v_i \in \mathcal{I}(T) \Leftrightarrow (\beta_1, ..., \beta_m) \in \mathcal{C}(A)$$

uma vez que

$$T(u) = T\left(\sum_{j=1}^{n} \alpha_j u_j\right) \underset{T \text{ é linear}}{=} \sum_{j=1}^{n} \alpha_j T(u_j) = \sum_{j=1}^{n} \alpha_j \sum_{i=1}^{m} a_{ij} v_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \alpha_j\right) v_i$$

e sendo $\{v_1, v_2, ..., v_m\}$ uma base de V tem-se

$$u \in \mathcal{N}(T) \Leftrightarrow T(u) = \mathbf{0} \Leftrightarrow \left(\sum_{j=1}^{n} a_{ij} \alpha_j = 0, \text{ para } i = 1, ..., m\right) \Leftrightarrow (\alpha_1, ..., \alpha_n) \in \mathcal{N}(A).$$

Além disso:

$$\mathcal{I}(T) = L\left(\{T(u_1), ..., T(u_n)\}\right) =$$
$$= L\left(\{a_{11}v_1 + ... + a_{m1}v_m, ..., a_{1n}v_1 + ... + a_{mn}v_m\}\right).$$

Teorema 73. Seja V um espaço linear de dimensão finita, com $\dim V = n$. Sejam $\mathcal{B}_1 = \{u_1, ..., u_n\}$ e $\mathcal{B}_2 = \{v_1, ..., v_n\}$ duas bases ordenadas de V. A representação matricial da transformação identidade $I : V \to V$ em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$ é igual à matriz de mudança da base $\mathcal{B}_1$ para $\mathcal{B}_2$. Isto é,

$$M(I; \mathcal{B}_1; \mathcal{B}_2) = S_{\mathcal{B}_1 \to \mathcal{B}_2}.$$

Teorema 74. Sejam U e V espaços lineares tais que $\dim U = n$ e $\dim V = m$. Seja $T : U \to V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ bases (ordenadas) de U e V respectivamente. Seja

$$A = M(T; \mathcal{B}_1; \mathcal{B}_2) \in \mathcal{M}_{m \times n}(\mathbb{R})$$

a matriz que representa T em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$. Tem-se então:

(i) $\dim \mathcal{N}(T) = \operatorname{nul} A$;

(ii) $\dim \mathcal{I}(T) = \operatorname{car} A$;

(iii) T é injectiva se e só se $\operatorname{nul} A = 0$, isto é, se e só se $\operatorname{car} A = n$;

(iv) T é sobrejectiva se e só se $\operatorname{car} A = m$.

Teorema 75. Sejam $\mathcal{B}_c^n = \{e_1, ..., e_n\}$ e $\mathcal{B}_c^m = \{e'_1, ..., e'_m\}$ as bases canónicas (ordenadas) de $\mathbb{R}^n$ e $\mathbb{R}^m$ respectivamente. Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ uma transformação linear. Considere-se a matriz

$$A = (a_{ij})_{m \times n} = M(T; \mathcal{B}_c^n; \mathcal{B}_c^m) \in \mathcal{M}_{m \times n}(\mathbb{R})$$

cuja coluna j, para cada $j = 1, ..., n$, é formada pelas coordenadas de $T(e_j)$ na base $\mathcal{B}_c^m$. Isto é,

$$T(e_j) = \sum_{i=1}^{m} a_{ij}e'_i = a_{1j} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + ... + a_{mj} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Então, tem-se, para todo o $u \in \mathbb{R}^n$,

$$T(u) = T(\lambda_1, ..., \lambda_n) = T\left(\sum_{j=1}^{n} \lambda_j e_j\right) = \sum_{j=1}^{n} \lambda_j T(e_j) = \sum_{j=1}^{n} \lambda_j \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} = Au.$$

Dem. Seja $u \in \mathbb{R}^n$. Então, existem $\lambda_1, ..., \lambda_n \in \mathbb{R}$ tais que

$$u = \lambda_1 e_1 + ... + \lambda_n e_n = \sum_{j=1}^{n} \lambda_j e_j.$$

Uma vez que, para todo o $j = 1, ..., n$, $T(e_j) = \sum_{i=1}^{m} a_{ij} e'_i$, tem-se

$$T(u) = T\left(\sum_{j=1}^{n} \lambda_j e_j\right) \underset{\text{T é linear}}{=} \sum_{j=1}^{n} \lambda_j T(e_j) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij} e'_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \lambda_j\right) e'_i =$$

$$= \left(\sum_{j=1}^{n} a_{1j} \lambda_j, ..., \sum_{j=1}^{n} a_{mj} \lambda_j\right) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = Au.$$

Observação 40. No caso em que $U = \mathbb{R}^n$, $V = \mathbb{R}^m$ e $\mathcal{B}_1 = \mathcal{B}_c^n$, $\mathcal{B}_2 = \mathcal{B}_c^m$, tem-se:

$$\mathcal{N}(T) = \mathcal{N}(A) \quad \text{e} \quad \mathcal{I}(T) = \mathcal{C}(A),$$

uma vez que neste caso as coordenadas de um vector numa base coincidem com o próprio vector.

Exemplo 52. **(i)** Seja $T : \mathbb{R}^4 \to \mathbb{R}^3$ definida por

$$T(x, y, z, w) = (3x + y - 2z, 0, x + 4z).$$

T é uma transformação linear e a matriz $M(T; \mathcal{B}_c^4; \mathcal{B}_c^3)$ que representa T em relação às bases canónicas (ordenadas) $\mathcal{B}_c^4$ e $\mathcal{B}_c^3$ de $\mathbb{R}^4$ e $\mathbb{R}^3$ respectivamente, é dada por

$$A = M(T; \mathcal{B}_c^4; \mathcal{B}_c^3) = \begin{bmatrix} 3 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 4 & 0 \end{bmatrix},$$

uma vez que $T(1, 0, 0, 0) = (3, 0, 1)$, $T(0, 1, 0, 0) = (1, 0, 0)$, $T(0, 0, 1, 0) = (-2, 0, 4)$ e $T(0, 0, 0, 1) = (0, 0, 0)$.

Tem-se então:

$$T(x, y, z, w) = M(T; \mathcal{B}_c^4; \mathcal{B}_c^3) \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = (3x + y - 2z, 0, x + 4z).$$

Além disso, tem-se

$$\begin{aligned} \mathcal{N}(T) &= \mathcal{N}(A) = \left\{(x, y, z, w) \in \mathbb{R}^4 : y = 14z \ \text{ e } \ x = -4z\right\} = \\ &= \{(-4z, 14z, z, w) : z, w \in \mathbb{R}\} = L(\{(-4, 14, 1, 0), (0, 0, 0, 1)\}) \end{aligned}$$

e

$$\mathcal{I}(T) = \mathcal{C}(A) = L(\{(3, 0, 1), (1, 0, 0)\}).$$

Uma base de $\mathcal{I}(T)$: $\{(3,0,1),(1,0,0)\}$. Uma base de $\mathcal{N}(T)$: $\{(-4,14,1,0),(0,0,0,1)\}$.

(ii) Sejam $\mathcal{B}_1 = \{1,t,t^2\}$ e $\mathcal{B}_2 = \{1,t,t^2,t^3\}$ as bases canónicas (ordenadas) de $\mathcal{P}_2$ e $\mathcal{P}_3$ respectivamente. Seja $D : \mathcal{P}_2 \to \mathcal{P}_3$ tal que $D(1) = 0$, $D(t) = 1$ e $D(t^2) = 2t$. D é uma transformação linear e a matriz $M(D;\mathcal{B}_1;\mathcal{B}_2)$ que representa D em relação às bases canónicas $\mathcal{B}_1$ e $\mathcal{B}_2$, é dada por

$$M(D;\mathcal{B}_1;\mathcal{B}_2) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Além disso tem-se

$$M(D;\mathcal{B}_1;\mathcal{B}_2)\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \\ 0 \end{bmatrix},$$

isto é, $D\left(a_0 + a_1 t + a_2 t^2\right) = a_1 + 2a_2 t$, com $a_0, a_1, a_2 \in \mathbb{R}$.

Além disso, como

$$\mathcal{N}(D) = \left\{a_0 + a_1 t + a_2 t^2 : D\left(a_0 + a_1 t + a_2 t^2\right) = \mathbf{0}\right\} = \left\{a_0 + a_1 t + a_2 t^2 : a_1 = a_2 = 0 \text{ e } a_0 \in \mathbb{R}\right\},$$

tem-se

$$\mathcal{N}(D) = \{a_0 : a_0 \in \mathbb{R}\} = L(\{1\}) \qquad \text{e} \qquad \mathcal{I}(D) = L(\{1, 2t\}).$$

Uma base de $\mathcal{I}(D)$: $\{1, 2t\}$. Uma base de $\mathcal{N}(D)$: $\{1\}$.

(iii) Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ a transformação linear cuja matriz que a representa em relação às bases ordenadas $\mathcal{B}_1 = \{(1,1,1),(0,1,1),(0,0,1)\}$ e $\mathcal{B}_2 = \{(1,1),(1,-1)\}$ de $\mathbb{R}^3$ e $\mathbb{R}^2$ respectivamente, é dada por

$$A = M(T;\mathcal{B}_1;\mathcal{B}_2) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}.$$

Seja $u \in \mathbb{R}^3$ e sejam $(\alpha_1, \alpha_2, \alpha_3)$ as coordenadas de u em relação à base $\mathcal{B}_1$. Tem-se

$$u \in \mathcal{N}(T) \Leftrightarrow (\alpha_1, \alpha_2, \alpha_3) \in \mathcal{N}(A)$$

e como

$$\mathcal{N}(A) = \mathcal{N}\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}\right) = \{(-2y-3z, y, z) : y, z \in \mathbb{R}\} = L(\{(-2,1,0),(-3,0,1)\}),$$

logo $\{(-2,1,0),(-3,0,1)\}$ é uma base de $\mathcal{N}(A)$ (uma vez que gera $\mathcal{N}(A)$ e é linearmente independente).

$$\begin{aligned} \mathcal{N}(T) &= \{(-2)(1,1,1) + 1(0,1,1) + 0(0,0,1), (-3)(1,1,1) + 0(0,1,1) + 1(0,0,1)\} = \\ &= L(\{(-2,-1,-1),(-3,-3,-2)\}). \end{aligned}$$

Logo $\{(-2,-1,-1),(-3,-3,-2)\}$ é uma base para $\mathcal{N}(T)$ (uma vez que gera $\mathcal{N}(T)$ e é linearmente independente).

Quanto ao contradomínio:

$$\mathcal{C}(A) = L\left(\{(1,2)\}\right),$$

logo $\{(1,2)\}$ é uma base de $\mathcal{C}(A)$ (uma vez que gera $\mathcal{C}(A)$ e é linearmente independente).

$$\mathcal{I}(T) = L\left(\{1(1,1)+2(1,-1)\}\right) = L\left(\{(3,-1)\}\right).$$

Uma base de $\mathcal{I}(T)$: $\{(3,-1)\}$ (uma vez que gera $\mathcal{I}(T)$ e é linearmente independente).

Note-se que:

$$\dim \mathcal{N}(T) = \dim \mathcal{N}(A) \qquad \dim \mathcal{I}(T) = \operatorname{car} A$$

e

$$\dim \mathcal{N}(T) + \dim \mathcal{I}(T) = \dim U \text{ (espaço de partida).}$$

Teorema 76. Sejam U, V e W espaços lineares de dimensões finitas. Sejam $\mathcal{B}_1, \mathcal{B}_2$ e $\mathcal{B}_3$ bases ordenadas de U, V e W respectivamente. Seja λ escalar. Sejam $T_1, T_2 \in \mathfrak{L}(U,V)$ e $T_3 \in \mathfrak{L}(V,W)$. Então, tem-se

$$M(T_1+\lambda T_2;\mathcal{B}_1;\mathcal{B}_2) = M(T_1;\mathcal{B}_1;\mathcal{B}_2) + \lambda M(T_2;\mathcal{B}_1;\mathcal{B}_2)$$

$$M(T_3\circ T_1;\mathcal{B}_1;\mathcal{B}_3) = M(T_3;\mathcal{B}_2;\mathcal{B}_3)M(T_1;\mathcal{B}_1;\mathcal{B}_2)$$

Dem. Se $A = M(T_1;\mathcal{B}_1;\mathcal{B}_2)$ e $B = M(T_2;\mathcal{B}_1;\mathcal{B}_2)$

$$\left(T_1+\lambda T_2\right)(u_j) = T_1(u_j) + \lambda T_2(u_j) = \sum_{i=1}^{m} a_{ij}v_i + \lambda \sum_{i=1}^{m} b_{ij}v_i = \sum_{i=1}^{m}\left(a_{ij}+\lambda b_{ij}\right)v_i$$

Logo

$$M(T_1+\lambda T_2;\mathcal{B}_1;\mathcal{B}_2) = A+\lambda B = M(T_1;\mathcal{B}_1;\mathcal{B}_2) + \lambda M(T_2;\mathcal{B}_1;\mathcal{B}_2).$$

Sejam agora $A = M(T_3;\mathcal{B}_2;\mathcal{B}_3)$ e $B = M(T_1;\mathcal{B}_1;\mathcal{B}_2)$

$$\left(T_3\circ T_1\right)(u_j) = T_3\left(T_1(u_j)\right) = T_3\left(\sum_{i=1}^{k} b_{ij}w_i\right) = \sum_{i=1}^{k} b_{ij}T_3(w_i) =$$

$$= \sum_{i=1}^{k} b_{ij}\sum_{l=1}^{m} a_{li}v_l = \sum_{l=1}^{m}\left(\underbrace{\sum_{i=1}^{k} a_{li}b_{ij}}_{\text{entrada } (l,j) \text{ de } AB}\right)v_l$$

Logo

$$M(T_3\circ T_1;\mathcal{B}_1;\mathcal{B}_3) = AB = M(T_3;\mathcal{B}_2;\mathcal{B}_3)M(T_1;\mathcal{B}_1;\mathcal{B}_2)$$

Teorema 77. Sejam U e V dois espaços lineares de dimensões finitas. Seja $T : U \to V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ duas bases ordenadas de U e V respectivamente. Seja $A = M(T;\mathcal{B}_1;\mathcal{B}_2)$ a matriz que representa T em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$.

Se $V = T(U)$ então T é invertível se e só se A fôr uma matriz quadrada invertível. Tem-se então

$$A^{-1} = M(T^{-1}; \mathcal{B}_2; \mathcal{B}_1),$$

isto é, A^{-1} será a matriz que representa T^{-1} em relação às bases $\mathcal{B}_2$ e $\mathcal{B}_1$.

Teorema 78. Sejam U e V espaços lineares de dimensões finitas respectivamente n e m. Isto é,

$$\dim U = n \quad \text{e} \quad \dim V = m.$$

Então, os espaços lineares $\mathfrak{L}(U, V)$ e $\mathcal{M}_{m\times n}(\mathbb{R})$ são isomorfos e escreve-se

$$\mathfrak{L}(U, V) \cong \mathcal{M}_{m\times n}(\mathbb{R}).$$

Tendo-se

$$\dim \mathfrak{L}(U, V) = mn.$$

Dem. Fixando bases ordenadas $\mathcal{B}_1$ e $\mathcal{B}_2$ para U e V respectivamente,

$$\begin{array}{ccc} \mathfrak{L}(U, V) & \rightarrow & \mathcal{M}_{m\times n}(\mathbb{R}) \\ T & \rightarrow & M(T; \mathcal{B}_1; \mathcal{B}_2) \end{array}$$

é uma transformação linear bijectiva.

Logo $\dim \mathfrak{L}(U, V) = \dim \mathcal{M}_{m\times n}(\mathbb{R}) = mn$

Teorema 79. Seja V um espaço linear de dimensão finita. Seja $T : V \rightarrow V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ duas bases ordenadas de V. Seja $M(T; \mathcal{B}_1; \mathcal{B}_1)$ a matriz que representa T em relação à base $\mathcal{B}_1$.

Então, a matriz $M(T; \mathcal{B}_2; \mathcal{B}_2)$ que representa T em relação à base $\mathcal{B}_2$, é dada por

$$M(T; \mathcal{B}_2; \mathcal{B}_2) = S_{\mathcal{B}_1\rightarrow\mathcal{B}_2} M(T; \mathcal{B}_1; \mathcal{B}_1) \left(S_{\mathcal{B}_1\rightarrow\mathcal{B}_2}\right)^{-1},$$

onde $S_{\mathcal{B}_1\rightarrow\mathcal{B}_2}$ é a matriz de mudança da base $\mathcal{B}_1$ para $\mathcal{B}_2$.

Isto é, o diagrama seguinte é comutativo.

$$\begin{array}{ccc} (V, \mathcal{B}_1) & \underset{T}{\overset{M(T;\mathcal{B}_1;\mathcal{B}_1)}{\longrightarrow}} & (V, \mathcal{B}_1) \\ S_{\mathcal{B}_1\rightarrow\mathcal{B}_2} \downarrow I & & I \downarrow S_{\mathcal{B}_1\rightarrow\mathcal{B}_2} \\ (V, \mathcal{B}_2) & \underset{M(T;\mathcal{B}_2;\mathcal{B}_2)}{\overset{T}{\longrightarrow}} & (V, \mathcal{B}_2) \end{array}$$

Note-se que neste caso:

$$T \circ I = I \circ T \Leftrightarrow$$

$$\Leftrightarrow M(T; \mathcal{B}_2; \mathcal{B}_2) S_{\mathcal{B}_1\rightarrow\mathcal{B}_2} = S_{\mathcal{B}_1\rightarrow\mathcal{B}_2} M(T; \mathcal{B}_1; \mathcal{B}_1) \Leftrightarrow$$

$$\Leftrightarrow M(T; \mathcal{B}_2; \mathcal{B}_2) = S_{\mathcal{B}_1\rightarrow\mathcal{B}_2} M(T; \mathcal{B}_1; \mathcal{B}_1) \left(S_{\mathcal{B}_1\rightarrow\mathcal{B}_2}\right)^{-1}.$$

Observação 41. Seja V um espaço linear de dimensão finita. Seja $T : V \rightarrow V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_2$ duas bases ordenadas de V. Seja $A = M(T; \mathcal{B}_1; \mathcal{B}_1)$

a matriz que representa T em relação à base $\mathcal{B}_1$ e seja $B = M(T; \mathcal{B}_2; \mathcal{B}_2)$ a matriz que representa T em relação à base $\mathcal{B}_2$. Como se tem

$$B = SAS^{-1}$$

então:

$$\operatorname{car} A = \operatorname{car} B \qquad \operatorname{nul} A = \operatorname{nul} B \qquad \operatorname{tr} A = \operatorname{tr} B \qquad \det A = \det B$$

Teorema 80. Caso geral. Sejam U e V dois espaços lineares de dimensões finitas. Seja $T : U \to V$ uma transformação linear. Sejam $\mathcal{B}_1$ e $\mathcal{B}_1'$ duas bases ordenadas de U. Sejam $\mathcal{B}_2$ e $\mathcal{B}_2'$ duas bases ordenadas de V. Seja $M(T; \mathcal{B}_1; \mathcal{B}_2)$ a matriz que representa T em relação às bases $\mathcal{B}_1$ e $\mathcal{B}_2$.

Então, a matriz $M(T; \mathcal{B}_1'; \mathcal{B}_2')$ que representa T em relação às bases $\mathcal{B}_1'$ e $\mathcal{B}_2'$, é dada por

$$M(T; \mathcal{B}_1'; \mathcal{B}_2') = S_{\mathcal{B}_2 \to \mathcal{B}_2'} M(T; \mathcal{B}_1; \mathcal{B}_2) \left(S_{\mathcal{B}_1 \to \mathcal{B}_1'}\right)^{-1},$$

onde $S_{\mathcal{B}_2 \to \mathcal{B}_2'}$ e $S_{\mathcal{B}_1 \to \mathcal{B}_1'}$ são as matrizes de mudança das bases $\mathcal{B}_2$ para $\mathcal{B}_2'$ e de $\mathcal{B}_1$ para $\mathcal{B}_1'$ respectivamente.

Isto é, o diagrama seguinte é comutativo.

$$\begin{array}{ccc}
(U, \mathcal{B}_1) & \overset{M(T;\mathcal{B}_1;\mathcal{B}_2)}{\underset{T}{\longrightarrow}} & (V, \mathcal{B}_2) \\
S_{\mathcal{B}_1 \to \mathcal{B}_1'} \downarrow I & & I \downarrow S_{\mathcal{B}_2 \to \mathcal{B}_2'} \\
(U, \mathcal{B}_1') & \underset{M(T;\mathcal{B}_1';\mathcal{B}_2')}{\overset{T}{\longrightarrow}} & (V, \mathcal{B}_2')
\end{array}$$

Note-se que neste caso:

$$T \circ I = I \circ T \Leftrightarrow$$

$$\Leftrightarrow M(T; \mathcal{B}_1'; \mathcal{B}_2') S_{\mathcal{B}_1 \to \mathcal{B}_1'} = S_{\mathcal{B}_2 \to \mathcal{B}_2'} M(T; \mathcal{B}_1; \mathcal{B}_2) \Leftrightarrow$$

$$\Leftrightarrow M(T; \mathcal{B}_1'; \mathcal{B}_2') = S_{\mathcal{B}_2 \to \mathcal{B}_2'} M(T; \mathcal{B}_1; \mathcal{B}_2) \left(S_{\mathcal{B}_1 \to \mathcal{B}_1'}\right)^{-1}.$$

Exemplo 53. Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$T(x, y) = (y, x, y - x).$$

T é uma transformação linear. A matriz $M(T; \mathcal{B}_c^2; \mathcal{B}_c^3)$ que representa T em relação à base canónica (ordenada) $\mathcal{B}_c^2$ de $\mathbb{R}^2$ e à base canónica (ordenada) $\mathcal{B}_c^3$ de $\mathbb{R}^3$, é dada por

$$M(T; \mathcal{B}_c^2; \mathcal{B}_c^3) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix}.$$

Sejam $\mathcal{B}_1 = \{(1, 1), (-1, 1)\}$ uma base ordenada de $\mathbb{R}^2$ e $\mathcal{B}_2 = \{(0, 0, 1), (0, 1, 1), (1, 1, 1)\}$ uma base ordenada de $\mathbb{R}^3$

A matriz $M(T; \mathcal{B}_1; \mathcal{B}_2)$ que representa T em relação à base ordenada $\mathcal{B}_1$ de $\mathbb{R}^2$ e à base ordenada $\mathcal{B}_2$ de $\mathbb{R}^3$, é dada por

$$M(T; \mathcal{B}_1; \mathcal{B}_2) = \begin{bmatrix} -1 & 3 \\ 0 & -2 \\ 1 & 1 \end{bmatrix},$$

uma vez que

$$\begin{aligned} T(1,1) &= (1,1,0) = -(0,0,1) + 0(0,1,1) + 1\,(1,1,1) \\ T(-1,1) &= (1,-1,2) = 3(0,0,1) - 2(0,1,1) + 1\,(1,1,1)\,. \end{aligned}$$

Vamos agora verificar que se tem

$$M(T;\mathcal{B}_1;\mathcal{B}_2) = S_{\mathcal{B}_c^3 \to \mathcal{B}_2} M(T;\mathcal{B}_c^2;\mathcal{B}_c^3) \left(S_{\mathcal{B}_c^2 \to \mathcal{B}_1}\right)^{-1}.$$

Uma vez que

$$\begin{gathered} (1,0,0) = 0(0,0,1) - 1(0,1,1) + 1\,(1,1,1)\,, \\ (0,1,0) = -(0,0,1) + 1(0,1,1) + 0\,(1,1,1)\,, \\ (0,0,1) = 1(0,0,1) + 0(0,1,1) + 0\,(1,1,1) \end{gathered}$$

tem-se então

$$S_{\mathcal{B}_c^3 \to \mathcal{B}_2} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Logo,

$$S_{\mathcal{B}_c^3 \to \mathcal{B}_2} M(T;\mathcal{B}_c^2;\mathcal{B}_c^3) \left(S_{\mathcal{B}_c^2 \to \mathcal{B}_1}\right)^{-1} = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ -1 & 1 \end{bmatrix} S_{\mathcal{B}_1 \to \mathcal{B}_c^2} =$$

$$= \begin{bmatrix} -2 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 0 & -2 \\ 1 & 1 \end{bmatrix} = M(T;\mathcal{B}_1;\mathcal{B}_2).$$

Por exemplo, para $(2,1) \in \mathbb{R}^2$, tem-se:

$$\begin{array}{ccc} \begin{array}{c}\text{coordenadas de } (2,1) \\ \text{na base } \mathcal{B}_c^2\end{array} & \xrightarrow[T]{M(T;\mathcal{B}_c^2;\mathcal{B}_c^3)} & \begin{array}{c}\text{coordenadas de } T(2,1) \\ \text{na base } \mathcal{B}_c^3\end{array} \\ & & \\ S_{\mathcal{B}_c^2 \to \mathcal{B}_1} \downarrow I & & I \downarrow S_{\mathcal{B}_c^3 \to \mathcal{B}_2} \\ & & \\ \begin{array}{c}\text{coordenadas de } (2,1) \\ \text{na base } \mathcal{B}_1\end{array} & \xrightarrow[M(T;\mathcal{B}_1;\mathcal{B}_2)]{T} & \begin{array}{c}\text{coordenadas de } T(2,1) \\ \text{na base } \mathcal{B}_2.\end{array} \end{array}$$

ou seja

$$\begin{array}{ccc} \begin{bmatrix} 2 \\ 1 \end{bmatrix} & \xrightarrow[T]{M(T;\mathcal{B}_c^2;\mathcal{B}_c^3)} & \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \\ & & \\ S_{\mathcal{B}_c^2 \to \mathcal{B}_1} \downarrow I & & I \downarrow S_{\mathcal{B}_c^3 \to \mathcal{B}_2} \\ & & \\ \begin{bmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{bmatrix} & \xrightarrow[M(T;\mathcal{B}_1;\mathcal{B}_2)]{T} & \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}. \end{array}$$

Valores próprios e vectores próprios de uma transformação linear. Diagonalização.

Definição 73. Seja V espaço linear. Seja $T : V \to V$ uma transformação linear. Diz-se que um escalar λ é um **valor próprio** de T se existir um vector não nulo $v \in V$ tal que

$$T(v) = \lambda v.$$

Aos vectores não nulos v que satisfaçam a equação anterior chamam-se **vectores próprios** associados ao valor próprio λ. Dado um valor próprio λ de T, o conjunto

$$E_\lambda = \{v \in V : T(v) = \lambda v\} = \mathcal{N}(T - \lambda I)$$

é um subespaço linear de V. Chama-se a E_λ o **subespaço próprio** associado ao valor próprio λ. À dimensão de E_λ chama-se **multiplicidade geométrica** de λ e denota-se por $m_g(\lambda)$, isto é,

$$\dim \mathcal{N}(T - \lambda I) = m_g(\lambda).$$

Exemplo 54. **(a)** Seja V um espaço linear e $I : V \to V$ a transformação identidade. Então todos os vectores de V, exceptuando o vector nulo, são vectores próprios de T associados ao valor próprio 1.

(b) Seja V o espaço linear das funções reais indefinidamente diferenciáveis em $\mathbb{R}$ e $T : V \to V$ a (transfomação) função derivada. Como, por exemplo

$$T\left(e^{2x}\right) = 2e^{2x}$$

então e^{2x} é vector próprio de T associado ao valor próprio 2.

Observação 42. **(i)** Sejam V um espaço linear e $\mathbf{0}$ o vector nulo de V. Seja $T : V \to V$ uma transformação linear. Um escalar λ é um valor próprio de T se e só se

$$\mathcal{N}(T - \lambda I) \neq \{\mathbf{0}\}.$$

(ii) Se o espaço linear V tiver dimensão finita n e se $A = M(T; \mathcal{B}; \mathcal{B})$ fôr a matriz $n \times n$ que representa T em relação a uma base ordenada $\mathcal{B}$ de V, então um escalar λ é um valor próprio de T se e só se esse escalar λ fôr solução da equação

$$\det(A - \lambda I) = 0,$$

uma vez que se tem, para $v \in V$,

$$(T - \lambda I)\, v = \mathbf{0} \Leftrightarrow$$

$$\Leftrightarrow (A - \lambda I) \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \mathbf{0}.$$

onde $\alpha_1, ..., \alpha_n$ são as coordenadas de v na base ordenada $\mathcal{B}$, daí que

$$\lambda \text{ é um valor próprio de } T \Leftrightarrow$$

$$\Leftrightarrow \mathcal{N}(T-\lambda I) \neq \{\mathbf{0}\} \Leftrightarrow$$

$$\Leftrightarrow \mathcal{N}(A-\lambda I) \neq \{\mathbf{0}\} \Leftrightarrow$$

$$\Leftrightarrow \det(A-\lambda I)=0$$

isto é

$$\lambda \text{ é um valor próprio de } T \Leftrightarrow \det(A-\lambda I)=0$$

Além disso, tem-se

$$v \text{ é um vector próprio de } T \Leftrightarrow v \in \mathcal{N}(T-\lambda I) \setminus \{\mathbf{0}\} \Leftrightarrow (\alpha_1, ..., \alpha_n) \in \mathcal{N}(A-\lambda I) \setminus \{\mathbf{0}\}$$

isto é

$$v \text{ é um vector próprio de } T \Leftrightarrow (\alpha_1, ..., \alpha_n) \in \mathcal{N}(A-\lambda I) \setminus \{\mathbf{0}\}$$

e

$$m_g(\lambda) = \dim \mathcal{N}(T-\lambda I) = \dim \mathcal{N}(A-\lambda I).$$

(iii) No caso em que $V=\mathbb{R}^n$ e $A=M(T;\mathcal{B}_c^n;\mathcal{B}_c^n)$, como (neste caso) $v=(\alpha_1,...,\alpha_n)$, tem-se

$$\mathcal{N}(T-\lambda I)=\mathcal{N}(A-\lambda I).$$

Definição 74. Seja U um espaço linear tal que $\dim U = n$. Seja $T: U \to U$ uma transformação linear. Diz-se que T **é diagonalizável** se existir uma base ordenada $\mathcal{B}$ de U em relação à qual a matriz $M(T;\mathcal{B};\mathcal{B})$ que representa T nessa base seja uma matriz diagonal.

Definição 75. Seja U um espaço linear de dimensão finita. Seja $T: U \to U$ uma transformação linear. Diz-se que T é uma **projecção** se

$$T \circ T = T \qquad (T^2 = T).$$

Teorema 81. Seja U um espaço linear de dimensão finita. Seja $T: U \to U$ uma transformação linear. Então:

(i) T projecção $\Rightarrow$ os valores próprios de T são 0 e 1

(ii) T projecção $\Rightarrow U = \mathcal{N}(T) \oplus \mathcal{I}(T)$

(iii) T projecção $\Rightarrow \mathcal{I}(T) = \mathcal{N}(T-I)$

(iv) T projecção $\Rightarrow U = \mathcal{N}(T) \oplus \mathcal{N}(T-I)$

(v) T projecção $\Rightarrow T$ é diagonalizável

Observação 43. Os valores próprios de T são 0 e 1 $\not\Rightarrow$ T projecção

Definição 76. Considere-se o produto interno usual. Seja U um espaço linear de dimensão finita. Seja $T: U \to U$ uma transformação linear. Diz-se que T é uma **projecção ortogonal** se

$$T \circ T = T \qquad (T^2 = T) \qquad \text{e} \qquad \mathcal{N}(T) = (\mathcal{I}(T))^{\perp}.$$

Diagonalização unitária e diagonalização ortogonal

Considera-se o produto interno usual.

Definição 77. Seja $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{C})$. Denota-se por A^H a matriz $\overline{A}^T$, isto é, a transposta da matriz conjugada $\overline{A} = (\overline{a_{ij}})$, onde $\overline{a_{ij}}$ é o complexo conjugado de a_{ij}. Ou seja, escreve-se $A^H = \overline{A}^T$. A matriz A diz-se **hermitiana** se

$$A^H = A.$$

Observação 44. **(a)** Sejam $\alpha, \beta \in \mathbb{C}$, $A, C \in \mathcal{M}_{m\times n}(\mathbb{C})$ e $B \in \mathcal{M}_{n\times r}(\mathbb{C})$. Tem-se:

(i) $\left(A^H\right)^H = A$ **(ii)** $(\alpha A + \beta B)^H = \overline{\alpha} A^H + \overline{\beta} B^H$ **(iii)** $(AC)^H = C^H A^H$

(b) Sendo A hermitiana tal que $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, então A é simétrica ($A^T = A$). Reciprocamente, se $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ fôr hermitiana então A é simétrica. Ou seja, para matrizes reais quadradas os conceitos de matriz simétrica e matriz hermitiana coincidem.

Teorema 82. Todos os valores próprios de uma matriz hermitiana são reais. Além disso, os vectores próprios associados a valores próprios distintos, de uma matriz hermitiana, são ortogonais.

Dem. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ tal que A é hermitiana. Seja λ um valor próprio de A e seja u um vector próprio associado. Seja $\alpha = u^H A u$. Então, tem-se

$$\overline{\alpha} = \alpha^H = \left(u^H A u\right)^H = u^H A^H \left(u^H\right)^H \underset{A \text{ é hermitiana}}{=} u^H A u = \alpha.$$

Ou seja, α é real. Por outro lado, como

$$\alpha = u^H A u = u^H \lambda u = \lambda \sum |u_i|^2 ,$$

tem-se

$$\lambda = \frac{\alpha}{\sum |u_i|^2} \in \mathbb{R}.$$

Sejam agora u_1 e u_2 vectores próprios associados respectivamente a valores próprios distintos λ_1 e λ_2. Então, tem-se

$$(Au_1)^H u_2 = u_1^H A^H u_2 \underset{A \text{ é hermitiana}}{=} u_1^H A u_2 = \lambda_2 u_1^H u_2$$

e

$$(Au_1)^H u_2 = (\lambda_1 u_1)^H u_2 = \overline{\lambda_1} u_1^H u_2 \underset{\lambda_1 \in \mathbb{R}}{=} \lambda_1 u_1^H u_2.$$

Logo, tem-se

$$\lambda_1 u_1^H u_2 = \lambda_2 u_1^H u_2 \Leftrightarrow (\lambda_1 - \lambda_2)\, u_1^H u_2 = 0.$$

E assim, como $\lambda_1 \neq \lambda_2$, então

$$\underbrace{u_1^H u_2}_{=\langle u_1, u_2\rangle} = 0,$$

ou seja, u_1 e u_2 são ortogonais.

Observação 45. Todos os valores próprios de uma matriz simétrica real são reais. Além disso, os vectores próprios associados a valores próprios distintos, de uma matriz simétrica, são ortogonais.

Definição 78. **(i)** Seja $U \in \mathcal{M}_{n\times n}(\mathbb{C})$. A matriz U diz-se **unitária** se se tiver $U^H U = I$, isto é, se $U^H = U^{-1}$, ou seja, se as colunas de U constituirem uma base ortonormada de $\mathbb{C}^n$.

(ii) Seja $P \in \mathcal{M}_{n\times n}(\mathbb{R})$. A matriz P diz-se **ortogonal** se se tiver $P^T P = I$, isto é, se $P^T = P^{-1}$, ou seja, se as colunas de P constituirem uma base ortonormada de $\mathbb{R}^n$.

Definição 79. **(i)** Seja $A \in \mathcal{M}_{n\times n}(\mathbb{C})$. A matriz A diz-se **unitariamente diagonalizável** se existir U^H unitária tal que UAU^H é uma matriz diagonal, isto é, se as colunas de U^H formarem uma base ortonormada de $\mathbb{C}^n$ constituída só por vectores próprios de A.

(ii) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. A matriz A diz-se **ortogonalmente diagonalizável** se existir P^T ortogonal tal que PAP^T é uma matriz diagonal, isto é, se as colunas de P^T formarem uma base ortonormada de $\mathbb{R}^n$ constituída só por vectores próprios de A.

Observação 46. **(i)** Seja U unitária tal que $U \in \mathcal{M}_{n\times n}(\mathbb{R})$. Então $U^H = U^T$, isto é, toda a matriz unitária real é ortogonal. Reciprocamente, se $P \in \mathcal{M}_{n\times n}(\mathbb{R})$ fôr ortogonal então P é unitária. Ou seja, para matrizes reais quadradas os conceitos de matriz ortogonal e matriz unitária coincidem.

(ii) Seja A uma matriz hermitiana. Se todos os valores próprios de A forem raízes simples do polinómio característico, então existe uma matriz unitária que diagonaliza A, isto é, existe U^H unitária tal que UAU^H é uma matriz diagonal, ou seja, A é unitariamente diagonalizável.

(iii) Como se vai ver a seguir, a afirmação anterior (ii) continua válida mesmo se os valores próprios não forem todos raízes simples do polinómio característico.

Teorema 83. (**Teorema de Schur**). Seja A uma matriz $n \times n$. Então, existe uma matriz unitária U^H tal que UAU^H é triangular superior (inferior).

Dem. A demonstração será efectuada por indução em n. O resultado é óbvio para $n = 1$. Suponhamos que a hipótese é válida para matrizes $k \times k$ e seja A uma matriz $(k+1)\times(k+1)$. Sejam λ_1 um valor próprio de A e w_1 um vector próprio associado de norma 1. Aplicando o método de ortogonalização de Gram-Schmidt, seja $\{w_1, \ldots, w_{k+1}\}$ uma base ortonormada para $\mathbb{C}^{k+1}$. Seja W^H a matriz cuja coluna i é igual ao vector w_i, para $i = 1, \ldots, k+1$. Então, por construção, a matriz W^H é unitária. Por outro lado, a primeira coluna de WAW^H é igual a WAw_1, tendo-se

$$WAw_1 = W\lambda_1 w_1 = \lambda_1 W w_1 = \lambda_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

e assim

$$WAW^H = \begin{bmatrix} \lambda_1 & | & * & \cdots & * \\ - & | & - & - & - \\ 0 & | & & & \\ \vdots & | & & M & \\ 0 & | & & & \end{bmatrix},$$

onde M é uma matriz $k \times k$.

Pela hipótese de indução, existe uma matriz $k \times k$ unitária $(V_1)^H$ tal que $V_1 M (V_1)^H = T_1$, onde T_1 é uma matriz triangular superior. Seja

$$V^H = \begin{bmatrix} 1 & | & 0 & \cdots & 0 \\ - & | & - & - & - \\ 0 & | & & & \\ \vdots & | & & (V_1)^H & \\ 0 & | & & & \end{bmatrix}.$$

Então V^H é unitária e tem-se

$$(VW)\, A\, (VW)^H = VWAW^H V^H =$$

$$= \begin{bmatrix} \lambda_1 & | & * & \cdots & * \\ - & | & - & - & - \\ 0 & | & & & \\ \vdots & | & & V_1 M (V_1)^H & \\ 0 & | & & & \end{bmatrix} = \begin{bmatrix} \lambda_1 & | & * & \cdots & * \\ - & | & - & - & - \\ 0 & | & & & \\ \vdots & | & & T_1 & \\ 0 & | & & & \end{bmatrix} = T,$$

onde T é uma matriz triangular superior. Como a matriz $(VW)^H$ é unitária, pondo $U^H = (VW)^H$, tem-se

$$UAU^H = T,$$

com T triangular superior e U^H unitária.

Exemplo 55. Seja $A = \begin{bmatrix} 2 & 1 \\ -2 & 5 \end{bmatrix}$. Os valores próprios de A são: 3 e 4. Como $\mathcal{N}(A - 3I) = L(\{(1,1)\})$ e $\mathcal{N}(A - 4I) = L(\{(1,2)\})$ então, aplicando Gram-Schmidt, o conjunto $\left\{ \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right), \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) \right\}$ é uma base ortonormada de $\mathbb{R}^2$ onde $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right)$ é um vector próprio de A associado ao valor próprio 3. Tem-se

$$UAU^H = T \quad \text{com} \quad U^H = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}.$$

Isto é

$$T = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 4 \end{bmatrix}.$$

Teorema 84. Seja A uma matriz hermitiana. Então existe uma matriz unitária U^H que diagonaliza A, isto é, A é unitariamente diagonalizável. Ou seja, existe U^H unitária tal que a matriz UAU^H é diagonal.

Dem. Pelo teorema anterior, existe uma matriz unitária U^H tal que a matriz UAU^H é triangular. Seja $T = UAU^H$. Tem-se então

$$T^H = \left(UAU^H\right)^H = \left(U^H\right)^H A^H U^H \underset{A \text{ é hermitiana}}{=} UAU^H = T.$$

Logo, como $T = T^H$ e T é triangular então T é diagonal.

Teorema 85. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica. Então existe uma matriz ortogonal P^T que diagonaliza A, isto é, A é ortogonalmente diagonalizável. Ou seja, existe P ortogonal tal que a matriz PAP^T é diagonal.

Observação 47. Sendo $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica, então existe P^T ortogonal tal que a matriz PAP^T é diagonal, isto é, existe uma base ortonormada de $\mathbb{R}^n$ formada só por vectores próprios de A, e a matriz P^T é a matriz cujas colunas são os vectores próprios de A que formam essa base ortonormada de $\mathbb{R}^n$, sendo PAP^T a matriz diagonal onde se coloca na entrada i da diagonal principal o valor próprio correspondente ao vector próprio da coluna i da matriz P^T.

Teorema 86. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. A é ortogonalmente diagonalizável $\Leftrightarrow$ A é simétrica

Dem. ($\Rightarrow$) Suponhamos que A é ortogonalmente diagonalizável. Sejam D diagonal e P^T ortogonal tais que $A = P^T DP$. Então

$$A^T = \left(P^T DP\right)^T = P^T D^T \left(P^T\right)^T = P^T DP = A.$$

($\Leftarrow$) Teorema anterior e o facto de todos os valores próprios de uma matriz simétrica real serem reais

Teorema 87. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica. Tem-se:

$$A \text{ é definida positiva, isto é, } u^T Au > 0 \text{ para todo o } u \neq \mathbf{0} \Leftrightarrow$$

$$\Leftrightarrow \text{todos os valores próprios de } A \text{ são positivos}$$

Dem. Sendo A simétrica então A é ortogonalmente diagonalizável, isto é, existem D diagonal e P^T ortogonal tais que $D = PAP^T$. Assim

$$(u^T Au > 0 \text{ para todo o } u \neq \mathbf{0}) \Leftrightarrow (\left(P^T u\right)^T AP^T u > 0 \text{ para todo o } u \neq \mathbf{0}) \Leftrightarrow$$

$$\Leftrightarrow (u^T \left(PAP^T\right) u > 0 \text{ para todo o } u \neq \mathbf{0}) \Leftrightarrow (u^T Du > 0 \text{ para todo o } u \neq \mathbf{0}) \Leftrightarrow$$

$$\Leftrightarrow (\sum_{i=1}^{n} (u_i)^2 \lambda_i > 0 \text{ para todo o } u \neq \mathbf{0}) \Leftrightarrow (\lambda_i > 0 \text{ para todo o } i = 1, ..., n)$$

onde $\lambda_1, ..., \lambda_n$ são os valores próprios de A são positivos.

Observação 48. **(i)** Existem matrizes não hermitianas que são unitariamente diagonalizáveis, como por exemplo as matrizes anti-hermitianas ($A^H = -A$) e as matrizes anti-simétricas ($A^T = -A$).

(ii) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{C})$. Suponhamos que A é **unitariamente diagonalizável**. Sejam D diagonal e U^H unitária tais que $A = U^H DU$. Como em geral se tem $D^H \neq D$, então

$$A^H = \left(U^H DU\right)^H = U^H D^H U \neq U^H DU = A.$$

Logo A não tem que ser necessariamente hermitiana.

(iii) O próximo teorema diz quais são as matrizes **unitariamente diagonalizáveis.**

Definição 80. Uma matriz A diz-se **normal** se

$$AA^H = A^H A.$$

Observação 49. Se $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ então A dir-se-á normal se

$$AA^T = A^T A.$$

Teorema 88. **(i)** Sendo $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ uma matriz normal tem-se para todo o vector u

$$\|Au\| = \left\|A^H u\right\|.$$

Em particular, sendo $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ uma matriz normal, para qualquer escalar λ, a matriz $A - \lambda I$ também é normal tendo-se

$$\|(A - \lambda I)\, u\| = \left\|(A - \lambda I)^H u\right\| = \left\|\left(A^H - \overline{\lambda} I\right) u\right\|$$

e assim, se λ fôr um valor próprio de A e u um vector próprio de A associado a esse valor próprio então $\overline{\lambda}$ é um valor próprio de A^H e u um vector próprio de A^H associado a esse valor próprio, isto é,

$$Au = \lambda u \quad \text{ e } \quad A^H u = \overline{\lambda} u.$$

(ii) Os vectores próprios associados a valores próprios distintos, de uma matriz normal, são ortogonais.

Dem. **(i)** Sendo $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ uma matriz normal tem-se para todo o vector u

$$\|Au\|^2 = (Au)^H Au = u^H A^H Au \underset{A^H A = AA^H}{=} u^H AA^H u = \left(A^H u\right)^H A^H u = \left\|A^H u\right\|^2$$

logo

$$\|Au\| = \left\|A^H u\right\|.$$

Sendo $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ uma matriz normal, para qualquer escalar λ, a matriz $A - \lambda I$ também é normal:

$$(A-\lambda I)(A-\lambda I)^H = (A-\lambda I)\left(A^H - \overline{\lambda} I\right) = AA^H - \lambda A^H - \overline{\lambda} A + |\lambda|^2 I = A^H A - \lambda A^H - \overline{\lambda} A + |\lambda|^2 I =$$

$$= A^H (A-\lambda I) - (A-\lambda I)\overline{\lambda} = \left(A^H - \overline{\lambda} I\right)(A-\lambda I) = (A-\lambda I)^H (A-\lambda I).$$

Logo

$$\|(A-\lambda I)u\| = \left\|(A-\lambda I)^H u\right\| = \left\|\left(A^H - \overline{\lambda} I\right)u\right\|$$

e assim, se λ fôr um valor próprio de A e u um vector próprio de A associado a esse valor próprio então $\overline{\lambda}$ é um valor próprio de A^H e u um vector próprio de A^H associado a esse valor próprio, isto é,

$$Au = \lambda u \qquad \text{e} \qquad A^H u = \overline{\lambda} u.$$

(ii) Seja $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ tal que A é normal. Sejam λ_1, λ_2 valores próprios de A tais que $\lambda_1 \neq \lambda_2$ e sejam v_1 e v_2 vectores próprios de A associados respectivamente a λ_1 e λ_2. Tem-se

$$Av_1 = \lambda_1 v_1 \qquad \text{e} \qquad A^H v_1 = \overline{\lambda_1} v_1$$

$$Av_2 = \lambda_2 v_2 \qquad \text{e} \qquad A^H v_2 = \overline{\lambda_2} v_2$$

e

$$(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = (\lambda_1 - \lambda_2)(v_1)^H v_2 = \lambda_1 (v_1)^H v_2 - \lambda_2 (v_1)^H v_2 =$$

$$= \left(\overline{\lambda_1} v_1\right)^H v_2 - (v_1)^H (\lambda_2 v_2) = \left(A^H v_1\right)^H v_2 - (v_1)^H (A v_2) = \left(A^H v_1\right)^H v_2 - \left(A^H v_1\right)^H v_2 = 0.$$

Assim, como $\lambda_1 \neq \lambda_2$, tem-se $\langle v_1, v_2\rangle = 0$.

Teorema 89. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{C})$. A é unitariamente diagonalizável $\Leftrightarrow$ A é normal

Dem. ($\Leftarrow$) Suponhamos que A é normal. Existe uma matriz unitária U^H e uma matriz triangular superior (inferior) T tais que $T = UAU^H$. Vejamos que T é normal. Tem-se

$$\begin{aligned} T^H T &= \left(UAU^H\right)^H UAU^H = UA^H U^H UAU^H = UA^H AU^H \underset{A \text{ é normal}}{=} \\ &= UAA^H U^H = UAU^H UA^H U^H = TT^H. \end{aligned}$$

Logo T é normal. Seja $T = (t_{ij})$ do tipo $n \times n$. Comparando as entradas das diagonais principais de TT^H e $T^H T$ tem-se:

$$\begin{aligned} |t_{11}|^2 + |t_{12}|^2 + |t_{13}|^2 + \cdots + |t_{1n}|^2 &= |t_{11}|^2 \\ |t_{22}|^2 + |t_{23}|^2 + \cdots + |t_{2n}|^2 &= |t_{12}|^2 + |t_{22}|^2 \\ &\vdots \\ |t_{nn}|^2 &= |t_{1n}|^2 + |t_{2n}|^2 + |t_{3n}|^2 + \cdots + |t_{nn}|^2 \end{aligned}$$

e assim, $t_{ij} = 0$ sempre que $i \neq j$. Logo T é diagonal e portanto A é unitariamente diagonalizável.

($\Rightarrow$) Suponhamos agora que A é unitariamente diagonalizável. Queremos mostrar que A é normal. Sejam D diagonal e U^H unitária tais que $D = UAU^H$, ou seja, $A = U^H DU$. Tem-se

$$AA^H = U^H DU \left(U^H DU\right)^H = U^H DUU^H D^H U = U^H \left(DD^H\right) U$$

e

$$A^H A = \left(U^H DU\right)^H U^H DU = U^H D^H UU^H DU = U^H \left(D^H D\right) U.$$

Como

$$DD^H = D^H D = \begin{bmatrix} |\lambda_1|^2 & 0 & \cdots & 0 \\ 0 & |\lambda_2|^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & |\lambda_n|^2 \end{bmatrix},$$

então tem-se $AA^H = A^H A$ e assim A é normal.

Exemplo 56. Seja

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

A matriz A não é simétrica logo não é ortogonalmente diagonalizável. Mas:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

isto é, $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ é normal e como tal é unitariamente diagonalizável. Tem-se

$$\underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{2} - \frac{1}{2}i\sqrt{3} & 0 \\ 0 & 0 & \frac{1}{2} + \frac{1}{2}i\sqrt{3} \end{bmatrix}}_{D} =$$

$$= \begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{6} + \frac{1}{2}i & -\frac{\sqrt{3}}{6} - \frac{1}{2}i & \frac{\sqrt{3}}{3} \\ -\frac{\sqrt{3}}{6} - \frac{1}{2}i & -\frac{\sqrt{3}}{6} + \frac{1}{2}i & \frac{\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} \frac{\sqrt{3}}{3} & -\frac{\sqrt{3}}{6} - \frac{1}{2}i & -\frac{\sqrt{3}}{6} + \frac{1}{2}i \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{3}}{6} + \frac{1}{2}i & -\frac{\sqrt{3}}{6} - \frac{1}{2}i \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix}}_{U^H}$$

onde 2, $\frac{1}{2} - \frac{1}{2}i\sqrt{3}$ e $\frac{1}{2} + \frac{1}{2}i\sqrt{3}$ são os valores próprios de A e

$$\left\{ \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right), \left(-\frac{\sqrt{3}}{6} - \frac{1}{2}i, -\frac{\sqrt{3}}{6} + \frac{1}{2}i, \frac{\sqrt{3}}{3}\right), \left(-\frac{\sqrt{3}}{6} + \frac{1}{2}i, -\frac{\sqrt{3}}{6} - \frac{1}{2}i, \frac{\sqrt{3}}{3}\right) \right\}$$

são respectivamente vectores próprios associados a esses valores próprios, normalizados e ortogonais entre si.

Raíz quadrada

Definição 81. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. À matriz B tal que $B^2 = A$ chama-se **raíz quadrada** de A e escreve-se

$$B = \sqrt{A}.$$

Exemplo 57. **(i)** Existem infinitas

$$\sqrt{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}},$$

por exemplo

$$\frac{1}{t}\begin{bmatrix} \mp s & \mp r \\ \mp r & \pm s \end{bmatrix}$$

com $s, r, t \in \mathbb{N}$ tais que $t^2 = s^2 + r^2$ (triplos pitagóricos).

(ii) Não existe $\sqrt{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}$.

Definição 82. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica e definida positiva. À matriz B tal que $B^2 = A$ chama-se **raíz quadrada** simétrica e definida positiva de A e escreve-se

$$B = \sqrt{A}.$$

Teorema 90. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica. Então, as seguintes afirmações são equivalentes.

(i) A é definida positiva.

(ii) Existe uma matriz simétrica definida positiva B tal que $A = B^2$. À matriz B chama-se **raíz quadrada** simétrica e definida positiva de A e escreve-se

$$B = \sqrt{A}.$$

(iii) Existe uma matriz invertível S tal que

$$A = S^T S.$$

Dem. **(i)** $\Rightarrow$ **(ii)** Supondo que A é definida positiva, vejamos que existe uma matriz simétrica definida positiva B tal que $A = B^2$.

Como A é simétrica, então A é ortogonalmente diagonalizável, isto é, existe uma matriz ortogonal P tal que

$$PAP^T = D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

onde $\lambda_1, ..., \lambda_n$ são os valores próprios de A, os quais são todos positivos por A ser definida positiva, tendo-se

$$D = (D')^2$$

com

$$D' = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & & 0 \\ 0 & \cdots & 0 & \sqrt{\lambda_n} \end{bmatrix}.$$

Assim

$$A = P^T D P = P^T (D')^2 P = \left(P^T D' P\right)\left(P^T D' P\right) = B^2$$

com

$$B = P^T D' P$$

simétrica:

$$B^T = \left(P^T D' P\right)^T = P^T (D')^T \left(P^T\right)^T = P^T D' P = B$$

e definida positiva uma vez que os valores próprios de $P^T D' P$ são os de D'.

(ii) $\Rightarrow$ **(iii)** Supondo que existe uma matriz simétrica definida positiva B tal que $A = B^2$, vejamos que existe uma matriz invertível S tal que

$$A = S^T S.$$

Como B é simétrica e definida positiva, basta fazer $S = B$ para ter-se

$$A = B^2 = BB = S^T S$$

com S simétrica e invertível uma vez que sendo B definida positiva, 0 não é valor próprio de B.

(iii) $\Rightarrow$ **(i)** Supondo que existe uma matriz invertível S tal que $A = S^T S$, vejamos que A é definida positiva, isto é, vejamos que

$$u^T A u > 0,$$

para todo o $u \neq \mathbf{0}$. Tem-se

$$u^T A u = u^T S^T S u = (Su)^T S u = \|Su\|^2 > 0$$

para todo o $u \neq \mathbf{0}$, uma vez que S é invertível.

Observação 50. **(i)** Sendo A matriz simétrica e definida positiva do tipo $n \times n$, existe uma única raíz quadrada simétrica e definida positiva B de A, isto é, existe uma única matriz B simétrica e definida positiva tal que $A = B^2$. No entanto, poderão existir pelo menos 2^n raízes quadradas de A, isto é, 2^n matrizes B para as quais se tem $A = B^2$.

(ii) Caso A seja simétrica e **semidefinida positiva** (todos os seus valores próprios são não negativos) então $\sqrt{A}$ também é única.

Exemplo 58. Seja $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$. Os valores próprios de A são: 3 e 5. Os vectores próprios associados ao valor próprio 3 são todos os vectores de $L\left(\{(-1,1)\}\right) \setminus \{\mathbf{0}\}$. Os vectores próprios associados ao valor próprio 5 são todos os vectores de $L\left(\{(1,1)\}\right) \setminus \{\mathbf{0}\}$. Tem-se

$$D = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = PAP^T$$

com

$$D = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} \quad \text{e} \quad P^T = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

onde

$$\left\{ \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right), \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) \right\}$$

são vectores próprios normalizados e ortogonais entre si respectivamente associados aos valores próprios 3 e 5. Logo

$$B = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{2}\sqrt{3} + \frac{1}{2}\sqrt{5} & \frac{1}{2}\sqrt{5} - \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{5} - \frac{1}{2}\sqrt{3} & \frac{1}{2}\sqrt{3} + \frac{1}{2}\sqrt{5} \end{bmatrix} = A^{1/2}$$

ou seja

$$B = \begin{bmatrix} \frac{1}{2}\sqrt{3} + \frac{1}{2}\sqrt{5} & \frac{1}{2}\sqrt{5} - \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{5} - \frac{1}{2}\sqrt{3} & \frac{1}{2}\sqrt{3} + \frac{1}{2}\sqrt{5} \end{bmatrix}$$

é a única matriz simétrica e definida positiva tal que

$$B^2 = \begin{bmatrix} \frac{1}{2}\sqrt{3} + \frac{1}{2}\sqrt{5} & \frac{1}{2}\sqrt{5} - \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{5} - \frac{1}{2}\sqrt{3} & \frac{1}{2}\sqrt{3} + \frac{1}{2}\sqrt{5} \end{bmatrix}^2 = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} = A.$$

Exemplo 59. Seja $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. Valores próprios de A: 0 e 3. Vejamos 2 modos para determinar $\sqrt{A}$:

(i) Base de $\mathbb{R}^3$ formada só por vectores próprios:

$$\Big\{ \underbrace{(-1,0,1)\,,(-1,1,0)}_{\in \mathcal{N}(A)\backslash\{\mathbf{0}\}}, \underbrace{(1,1,1)}_{\in \mathcal{N}(A-3I)\backslash\{\mathbf{0}\}} \Big\}.$$

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} -1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}}_{P^{-1}} \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}}_{P} =$$

$$= \left(\underbrace{\begin{bmatrix} -1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}}_{P^{-1}} \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix}}_{\sqrt{D}} \underbrace{\begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}}_{P} \right)^2$$

$$\sqrt{A} = \underbrace{\begin{bmatrix} -1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}}_{P^{-1}} \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix}}_{\sqrt{D}} \underbrace{\begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}}_{P} = \begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix}$$

(ii) Base ortogonal de $\mathbb{R}^3$ formada só por vectores próprios:

$$\left\{ \underbrace{(-1,0,1)\,,(-1,1,0) - \frac{1}{2}(-1,0,1)}_{\in \mathcal{N}(A)\backslash\{\mathbf{0}\}}, \underbrace{(1,1,1)}_{\in \mathcal{N}(A-3I)\backslash\{\mathbf{0}\}} \right\} = \{(-1,0,1)\,,(-1,2,-1)\,,(1,1,1)\}\,.$$

Base ortonormada de $\mathbb{R}^3$ formada só por vectores próprios:

$$\left\{ \underbrace{\left(-\frac{\sqrt{2}}{2},0,\frac{\sqrt{2}}{2}\right), \left(-\frac{\sqrt{6}}{6},\frac{\sqrt{6}}{3},-\frac{\sqrt{6}}{6}\right)}_{\in \mathcal{N}(A)\backslash\{\mathbf{0}\}}, \underbrace{\left(\frac{\sqrt{3}}{3},\frac{\sqrt{3}}{3},\frac{\sqrt{3}}{3}\right)}_{\in \mathcal{N}(A-3I)\backslash\{\mathbf{0}\}} \right\}.$$

$$\sqrt{A} = \underbrace{\begin{bmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \end{bmatrix}}_{P^T} \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix}}_{\sqrt{D}} \underbrace{\begin{bmatrix} -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{3} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix}}_{P} == \begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \end{bmatrix}.$$

Mínimos quadrados

Existem aplicações relativamente às quais os erros cometidos nas medições das entradas de A ou de b podem levar a que o sistema de equações lineares $Au = b$ não tenha solução, quando teoricamente deveria ter. Em tais casos é natural a procura da "melhor solução aproximada" para esse problema.

Considera-se o produto interno usual.

Definição 86. Sejam $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e $b \in \mathbb{R}^m$. Então, a $u \in \mathbb{R}^n$ chama-se melhor solução aproximada ou **solução de mínimos quadrados** de $Au = b$ se u satisfizer

$$Au = P_{\mathcal{C}(A)}(b)$$

para qualquer $u \in \mathbb{R}^n$, onde $P_{\mathcal{C}(A)}$ é a projecção ortogonal de $\mathbb{R}^m$ sobre $\mathcal{C}(A)$. Ao vector $b - Au$ chama-se vector erro de mínimos quadrados e ao escalar $\|b - Au\|$ chama-se erro de mínimos quadrados.

Observação 51. Sejam $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e $b \in \mathbb{R}^m$. Procuremos então um método para determinar as soluções de mínimos quadrados de $Au = b$. Atendendo a que $Au \in \mathcal{C}(A)$ para todo o $u \in \mathbb{R}^n$, tem-se

$$\left\|b - P_{\mathcal{C}(A)}(b)\right\| \leq \|b - Au\|,$$

para todo o $u \in \mathbb{R}^n$, sendo a distância $\|b - Au\|$ é mínima se

$$Au = P_{\mathcal{C}(A)}(b),$$

Como $P_{\mathcal{C}(A)}(b) \in \mathcal{C}(A)$, a equação $Au = P_{\mathcal{C}(A)}(b)$ tem sempre solução e essas soluções são as soluções de mínimos quadrados de $Au = b$. Deste modo, qualquer sistema de equações lineares tem sempre pelo menos uma solução de mínimos quadrados.

Por outro lado, tem-se

$$Au = P_{\mathcal{C}(A)}(b) \Leftrightarrow P_{\mathcal{C}(A)}Au = P_{\mathcal{C}(A)}b \Leftrightarrow P_{\mathcal{C}(A)}(Au - b) = \mathbf{0} \Leftrightarrow$$

$$\Leftrightarrow Au - b \in \mathcal{N}\left(P_{\mathcal{C}(A)}\right) = (\mathcal{C}(A))^{\perp} = \mathcal{N}\left(A^T\right) \Leftrightarrow A^T(Au - b) = \mathbf{0} \Leftrightarrow A^TAu = A^Tb.$$

À equação

$$A^TAu = A^Tb.$$

chama-se equação normal associada a $Au = b$.

Teorema 93. Sejam $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ e $b \in \mathbb{R}^m$.

(i) As soluções de mínimos quadrados do sistema de equações lineares

$$Au = b$$

são as soluções da equação normal

$$A^TAu = A^Tb.$$

(ii) Se $\operatorname{car} A = n$ então a equação normal

$$A^T A u = A^T b$$

tem a solução única

$$u = \left(A^T A\right)^{-1} A^T b$$

e tem-se

$$P_{\mathcal{C}(A)}\left(b\right) = Au = A\left(A^T A\right)^{-1} A^T b,$$

isto é,

$$A\left(A^T A\right)^{-1} A^T$$

é a matriz que representa a projecção ortogonal $P_{\mathcal{C}(A)}$ na base canónica $\mathcal{B}_c^m$ de $\mathbb{R}^m$, isto é,

$$A\left(A^T A\right)^{-1} A^T = M\left(P_{\mathcal{C}(A)}; \mathcal{B}_c^m; \mathcal{B}_c^m\right).$$

Teorema 94. Seja $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Então

$$\operatorname{car} A = \operatorname{car}\left(A^T A\right).$$

Dem. Basta para isso, mostrar que

$$\mathcal{N}\left(A\right) = \mathcal{N}\left(A^T A\right).$$

Seja $u \in \mathcal{N}\left(A\right)$. Como $Au = \mathbf{0}$ então $A^T A u = A^T \mathbf{0} = \mathbf{0}$ e assim $u \in \mathcal{N}\left(A^T A\right)$.
Reciprocamente, seja $u \in \mathcal{N}\left(A^T A\right)$ e vejamos que $u \in \mathcal{N}\left(A\right)$. Tem-se $A^T A u = \mathbf{0}$, logo

$$Au \in \mathcal{N}\left(A^T\right) = \left(\mathcal{L}\left(A^T\right)\right)^{\perp} = \left(\mathcal{C}\left(A\right)\right)^{\perp}$$

e como tal

$$\langle Au, Au \rangle = 0,$$

ou seja $\left\|Au\right\|^2 = 0$ e então $Au = \mathbf{0}$, isto é, $u \in \mathcal{N}\left(A\right)$.

Observação 52. Vejamos agora o modo como se pode determinar uma curva (ou recta) específica que se possa "ajustar" a um conjunto de pontos determinados experimentalmente.

(i) A partir de dois ou mais pontos dados

$$\left(x_1, y_1\right), \left(x_2, y_2\right), \ldots, \left(x_m, y_m\right),$$

pretende-se determinar uma recta $y = a_0 + a_1 x$ que seja a recta que "melhor aproxime" ou a recta de mínimos quadrados de melhor ajuste aos pontos dados (recta de regressão). Isto é, pretende-se determinar as soluções de mínimos quadrados de

$$\begin{cases} y_1 = a_0 + a_1 x_1 \\ y_2 = a_0 + a_1 x_2 \\ \quad\vdots \\ y_m = a_0 + a_1 x_m \end{cases}$$

ou seja de

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Sejam

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}, \qquad u = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \qquad \text{e} \qquad b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Atendendo a que $\text{car } A = \text{car}\left(A^T A\right)$, se houver pelo menos dois pontos distintos, tem-se $\text{car } A = 2$ e nesse caso, a equação normal

$$A^T A u = A^T b$$

tem como única solução de mínimos quadrados

$$u = \left(A^T A\right)^{-1} A^T b.$$

Assim, a recta de mínimos quadrados $y = a + bx$ é a recta que torna mínimos os quadrados cuja soma

$$(y_1 - (a_0 + a_1 x_1))^2 + (y_2 - (a_0 + a_1 x_2))^2 + \cdots + (y_m - (a_0 + a_1 x_m))^2$$

é dada por

$$\|b - Au\|^2,$$

onde $\|b - Au\|$ é o erro de mínimos quadrados.

(ii) A partir de m pontos dados $(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)$, pretende-se determinar um polinómio cujo gráfico esteja tão perto quanto possível desses m pontos dados. Isto é, com $m \in \mathbb{N}$ previamente fixo, pretende-se determinar as soluções de mínimos quadrados do sistema de m equações a $n + 1$ incógnitas $(a_0, a_1, a_2, \ldots, a_n)$

$$\begin{cases} y_1 = a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n \\ y_2 = a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_n x_2^n \\ \vdots \\ y_m = a_0 + a_1 x_m + a_2 x_m^2 + \cdots + a_n x_m^n \end{cases}$$

ou seja de

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Sejam

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix}, \qquad u = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \qquad \text{e} \qquad b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}.$$

Note-se que se $n+1=m$ e se os pontos dados forem distintos, então existe um único polinómio de grau n (o chamado polinómio interpolador) que passa por todos esses m pontos.

Por outro lado, atendendo a que $\operatorname{car} A = \operatorname{car}\left(A^T A\right)$, se $n<m$ e pelo menos $n+1$ pontos forem distintos, tem-se $\operatorname{car} A = n+1$ e então a equação normal

$$A^T A u = A^T b$$

tem como única solução de mínimos quadrados

$$u = \left(A^T A\right)^{-1} A^T b.$$

Exemplo 61. Determinemos a recta de mínimos quadrados relativa aos pontos

$$(0,1)\,,(1,3)\,,(2,4) \text{ e } (3,4)\,.$$

Sejam

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \qquad \text{e} \qquad b = \begin{bmatrix} 1 \\ 3 \\ 4 \\ 4 \end{bmatrix}.$$

Tem-se $\operatorname{car} A = 2$ e como tal a solução de mínimos quadrados é única e dada por:

$$u = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \left(A^T A\right)^{-1} A^T b =$$

$$= \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 1 \end{bmatrix},$$

tendo-se

$$y = \frac{3}{2} + x.$$

O vector $b - Au$ é o vector erro de mínimos quadrados, sendo o erro de mínimos quadrados dado por:

$$\|b - Au\| =$$

$$= \sqrt{\left(y_1 - (a_0 + a_1 x_1)\right)^2 + \left(y_2 - (a_0 + a_1 x_2)\right)^2 + \left(y_3 - (a_0 + a_1 x_3)\right)^2 + \left(y_4 - (a_0 + a_1 x_4)\right)^2} =$$

$$= \sqrt{\left(1 - \left(\frac{3}{2} + 0\right)\right)^2 + \left(3 - \left(\frac{3}{2} + 1\right)\right)^2 + \left(4 - \left(\frac{3}{2} + 2\right)\right)^2 + \left(4 - \left(\frac{3}{2} + 3\right)\right)^2} =$$

$$= \sqrt{\frac{25}{16} + \frac{1}{4} + \frac{1}{16} + \frac{1}{4}} = \frac{\sqrt{34}}{4}.$$

Matrizes elementares e factorização triangular

Definição 84. Uma **matriz elementar** é uma matriz do tipo $n \times n$ obtida da matriz identidade I (do tipo $n \times n$) através de uma única operação elementar.

(i) A matriz P_{ij}, chamada **matriz de permutação**, é a matriz elementar obtida por troca da linha i com a linha j da matriz I. Tem-se:

$$P_{ij} = \left[\begin{array}{ccccccccccc}
1 & 0 & \cdots & & & & & & & \cdots & 0 \\
0 & \ddots & \ddots & & & & & & & & \vdots \\
\vdots & \ddots & 1 & & & & & & & & \\
 & & & 0 & & & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & & & 0 & & & \\
 & & & & & & & & 1 & \ddots & \vdots \\
\vdots & & & & & & & & \ddots & \ddots & 0 \\
0 & \cdots & & & & & & & \cdots & 0 & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\
\end{array}.$$

(ii) A matriz $E_i(\alpha)$ é a matriz elementar obtida da matriz I através do produto do escalar $\alpha \neq 0$ pela linha i da matriz I. Tem-se:

$$E_i(\alpha) = \left[\begin{array}{ccccccc}
1 & 0 & \cdots & & & \cdots & 0 \\
0 & \ddots & \ddots & & & & \vdots \\
\vdots & \ddots & 1 & & & & \\
 & & & \alpha & & & \\
 & & & & 1 & \ddots & \vdots \\
\vdots & & & & \ddots & \ddots & 0 \\
0 & \cdots & & & \cdots & 0 & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \\ \leftarrow i \\ \\ \\ \\
\end{array}.$$

(iii) A matriz $E_{ij}(\alpha)$ é a matriz elementar obtida da matriz I por soma da linha j com um múltiplo escalar α da linha i. Por exemplo para $i < j$ tem-se:

$$E_{ij}(\alpha) = \left[\begin{array}{ccccccc}
1 & 0 & \cdots & & & \cdots & 0 \\
0 & \ddots & \ddots & & & & \vdots \\
\vdots & \ddots & 1 & & & & \\
 & & & \ddots & & & \\
 & & \alpha & & 1 & \ddots & \vdots \\
\vdots & & & & \ddots & \ddots & 0 \\
0 & \cdots & & & \cdots & 0 & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\
\end{array}.$$

Observação 53. **(i)** As matrizes elementares $E_{ij}(\alpha)$, com $i < j$, são matrizes triangulares inferiores.

(ii) As matrizes elementares $E_{ij}(\alpha)$ e $E_{ik}(\beta)$ comutam, isto é, $E_{ij}(\alpha)E_{ik}(\beta) = E_{ik}(\beta)E_{ij}(\alpha)$.

Exemplo 61. Sejam α, β escalares com $\alpha \neq 0$. As matrizes elementares do tipo 2×2 são:

$$P_{12} = P_{21} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ E_1(\alpha) = \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix},\ E_2(\alpha) = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix},\ E_{12}(\beta) = \begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix} \text{ e } E_{21}(\beta) = \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix}.$$

Observação 54. Sejam E uma matriz elementar do tipo $m \times m$ e A uma matriz qualquer do tipo $m \times n$. Então, EA é a matriz obtida de A através da mesma operação elementar que originou E. Isto é, aplicar uma operação elementar a uma matriz corresponde a multiplicar essa matriz à esquerda por uma matriz elementar.

Exemplo 62. Considere-se a matriz aumentada $\left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right]$. A operação elementar:

$$\left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right] \underset{L_1 \leftrightarrow L_3}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 5 & 15 & -10 & 40 & -45 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right],$$

corresponde à seguinte multiplicação (à esquerda):

$$\underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}}_{P_{13}} \left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 5 & 15 & -10 & 40 & -45 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right].$$

A operação elementar:

$$\left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 5 & 15 & -10 & 40 & -45 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{\frac{1}{5}L_2 \to L_2}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right],$$

corresponde à seguinte multiplicação (à esquerda):

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E_2(\frac{1}{5})} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 5 & 15 & -10 & 40 & -45 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right].$$

A operação elementar:

$$\left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{-L_1+L_2 \to L_2}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right],$$

corresponde à seguinte multiplicação (à esquerda):

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E_{12}(-1)} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 1 & 3 & -2 & 8 & -9 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right].$$

Finalmente, a operação elementar:

$$\left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] \underset{3L_2+L_3 \to L_3}{\longrightarrow} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

corresponde à seguinte multiplicação (à esquerda):

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}}_{E_{23}(3)} \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 3 & -9 & 6 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Tem-se então:

$$E_{23}\,(3)\,E_{12}\,(-1)\,E_2\left(\frac{1}{5}\right)P_{13} \left[\begin{array}{cccc|c} 0 & 0 & 3 & -9 & 6 \\ 5 & 15 & -10 & 40 & -45 \\ 1 & 3 & -1 & 5 & -7 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 3 & -1 & 5 & -7 \\ 0 & 0 & -1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Observação 55. Toda a matriz elementar é invertível e a respectiva inversa é também uma matriz elementar. Tem-se:

(i) $(P_{ij})^{-1} = P_{ij}$. **(ii)** $(E_i(\alpha))^{-1} = E_i(1/\alpha)$, para $\alpha \neq 0$. **(iii)** $(E_{ij}(\alpha))^{-1} = E_{ij}(-\alpha)$.

Observação 56. Uma matriz A é invertível se e só se fôr igual ao produto de matrizes elementares. Tem-se assim um modo alternativo para calcular a matriz inversa de uma matriz invertível.

Observação 57. **(i)** O produto de duas matrizes triangulares inferiores (superiores) é uma matriz triangular inferior (superior).

(ii) Se uma matriz triangular superior (inferior) fôr invertível então a sua inversa é também triangular superior (inferior).

Dem. **(i)** Se $A = (a_{ij})$ e $B = (b_{ij})$ com $a_{ij} = b_{ij} = 0$ se $i < j$; então para $AB = \left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)$ tem-se $\sum_{k=1}^{n} a_{ik}b_{kj} = 0$ se $i < j$ uma vez que se $k \geqslant j > i$ então $a_{ik} = 0$ e se $j > k$ então $b_{kj} = 0$. Além disso a diagonal principal da matriz triangular inferior AB é dada por: $\sum_{k=1}^{n} a_{ik}b_{ki} = a_{ii}b_{ii}$, $i = 1, ..., n$, uma vez que se $i < k$ então $a_{ik} = 0$ e se $k < i$ então $b_{ki} = 0$.

Teorema 92. Seja A uma matriz do tipo $m \times n$. Então ou A admite a factorização $A = LU$ ou existe uma matriz de permutação P tal que PA admite a factorização $PA = LU$, onde L é uma matriz triangular inferior com as entradas da diagonal principal todas iguais a 1 e U é uma matriz em escada.

Teorema 93. Seja A do tipo $n \times n$ uma matriz invertível. Então ou A admite a factorização **única** $A = LU$ ou existe uma matriz de permutação P tal que PA admite a factorização **única** $PA = LU$, onde L é uma matriz triangular inferior com as entradas da diagonal principal todas iguais a 1 e U é uma matriz triangular superior cujas entradas da diagonal principal são os **pivots** que resultam de aplicar o método de eliminação de Gauss à matriz A.

Exemplo 63. Seja $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 4 \\ 2 & 3 & 5 \end{bmatrix}$. Tem-se $E_{23}(1)E_{13}(-2)E_{12}(-2)A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix}$. Logo,

$$A = (E_{12}(-2))^{-1}\,(E_{13}(-2))^{-1}\,(E_{23}(1))^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix}.$$

Isto é,

$$A = E_{12}(2)E_{13}(2)E_{23}(-1) \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix}, \qquad \text{ou ainda,} \qquad A = LU,$$

com

$$L = E_{12}(2)E_{13}(2)E_{23}(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \qquad \text{e} \qquad U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 5 \end{bmatrix}.$$

Exemplo 64. Seja $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 10 & 6 \\ 0 & 1 & 7 & 8 \end{bmatrix}$. Tem-se $P_{24}A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 7 & 8 \\ 0 & 0 & 10 & 6 \\ 0 & 0 & 5 & 6 \end{bmatrix}$ e

$$E_{34}\left(-1/2\right) P_{24}A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 7 & 8 \\ 0 & 0 & 10 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix}. \qquad \text{Logo} \qquad P_{24}A = E_{34}\left(1/2\right) \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 7 & 8 \\ 0 & 0 & 10 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix},$$

ou ainda,

$$PA = LU, \quad \text{com} \quad P = P_{24}, \qquad L = E_{34}\left(1/2\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 1 \end{bmatrix} \qquad \text{e} \quad U = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 7 & 8 \\ 0 & 0 & 10 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Observação 58. O determinante de cada um dos três tipos de matrizes elementares é dado por

$$\det P_{ij} = -1, \qquad \det E_i\left(\alpha\right) = \alpha, \qquad \det E_{ij}\left(\alpha\right) = 1.$$

Assim $\det\left(EA\right) = \det E \det A$, onde E é uma matriz elementar ($P_{ij}, E_i\left(\alpha\right)$ ou $E_{ij}\left(\alpha\right)$). O que conduz à demonstração de se ter

$$\det\left(AB\right) = \det A \det B.$$

Teorema 94. Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. As seguintes afirmações são equivalentes.

(i) A é igual ao produto de matrizes elementares.

(ii) A é invertível.

(iii) $A^T A$ é invertível.

(iv) $\operatorname{nul} A = 0$.

(v) $\operatorname{car} A = n$.

(vi) $Au = \mathbf{0}$ tem apenas a solução trivial $u = \mathbf{0}$.

(vii) $Au = b$ tem solução única u para cada $b \in \mathbb{R}^n$.

(viii) $\det A \neq 0$.

(ix) $\mathcal{N}(A) = \{\mathbf{0}\}$.

(x) As colunas de A geram $\mathbb{R}^n$.

(xi) As colunas de A são independentes.

(xii) As colunas de A formam uma base de $\mathbb{R}^n$.

(xiii) As linhas de A geram $\mathbb{R}^n$.

(xiv) As linhas de A são independentes.

(xv) As linhas de A formam uma base de $\mathbb{R}^n$.

(xvi) A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(u) = Au$, para $u \in \mathbb{R}^n$, é sobrejectiva.

(xvii) A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(u) = Au$, para $u \in \mathbb{R}^n$, é injectiva.

(xviii) A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(u) = Au$, para $u \in \mathbb{R}^n$, é bijectiva.

(xix) A transformação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(u) = Au$, para $u \in \mathbb{R}^n$, é invertível.

(xx) 0 não é valor próprio de A.

(xxi) $(\mathcal{N}(A))^{\perp} = \mathbb{R}^n$.

(xxii) $(\mathcal{L}(A))^{\perp} = \{\mathbf{0}\}$.

Formas quadráticas

Considera-se o produto interno usual.

Definição 85. Uma **equação quadrática** em duas variáveis x_1 e x_2 é uma equação da forma

$$ax_1^2 + bx_2^2 + 2cx_1x_2 + dx_1 + ex_2 + f = 0$$

a qual pode ser escrita na forma

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & c \\ c & b \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + f = 0.$$

Sejam

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \text{e} \qquad A = \begin{bmatrix} a & c \\ c & b \end{bmatrix}.$$

(A é uma matriz real simétrica). À função real a duas variáveis reais $Q : \mathbb{R}^2 \to \mathbb{R}$ definida por $Q(x) = x^T Ax$, com

$$x^T Ax = ax_1^2 + bx_2^2 + 2cx_1x_2$$

chama-se **forma quadrática** real a 2 variáveis reais associada à equação quadrática anterior.

Podem haver equações do 2^o grau e formas quadráticas com um n^o de variáveis superior a 2. Uma equação quadrática em n variáveis $x_1, x_2, \ldots, x_n$ é uma equação da forma

$$x^T Ax + Bx + \alpha = 0,$$

onde $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $A = (a_{ij})$ é uma matriz real simétrica do tipo $n \times n$, $B \in \mathcal{M}_{1\times n}(\mathbb{R})$ e α é um escalar. À função real a n variáveis reais $Q : \mathbb{R}^n \to \mathbb{R}$ definida por

$$Q(x) = x^T Ax = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij}x_j \right) x_i$$

chama-se **forma quadrática** real a n variáveis reais associada à equação quadrática anterior. Se a matriz A não fôr simétrica considera-se

$$B = \frac{A + A^T}{2}$$

e tem-se a mesma expressão, isto é

$$x^T Bx = x^T \frac{A + A^T}{2} x = \frac{1}{2} x^T Ax + \frac{1}{2} x^T A^T x = \frac{1}{2} x^T Ax + \frac{1}{2} x^T Ax = x^T Ax$$

uma vez que tendo-se $x^T Ax \in \mathbb{R}$ obtém-se

$$x^T Ax = \left(x^T Ax\right)^T = x^T A^T x.$$

Teorema 95. (Teorema dos eixos principais). Seja $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que A é simétrica. Então existe uma mudança de variáveis ortogonal que transforma a forma quadrática $x^T Ax$ na forma quadrática $y^T Dy$ sem termos cruzados. Isto é, se P^T diagonalizar A ortogonalmente ($D = PAP^T$), então a mudança de variáveis $x = P^T y$ transforma a forma quadrática $x^T Ax$ na forma quadrática $y^T Dy$:

$$x^T Ax = y^T PAP^T y = y^T Dy = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 =$$

$$= \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

onde $\lambda_1, \lambda_2, \ldots, \lambda_n$ são os valores próprios de A associados respectivamente aos vectores próprios que constituem as colunas de P^T e que formam uma base ortonormada de $\mathbb{R}^n$.

Definição 86. (i) Chama-se **cónica** ou secção cónica à curva plana obtida por meio de um corte efectuado por um plano relativamente a uma superfície cónica. As secções cónicas que se obtêm quando o plano que efectua o corte não passa pelo vértice da superfície cónica, são elipses (os valores próprios têm o mesmo sinal) (podendo ter-se circunferências: quando o corte é efectuado perpendicularmente ao eixo de simetria do cone), parábolas (um dos dois valores próprios é zero) e hipérboles (os dois valores próprios têm sinais contrários).

(ii) Em $\mathbb{R}^3$ tem-se

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad A = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix} \qquad \text{e} \qquad B = \begin{bmatrix} g \\ h \\ i \end{bmatrix}$$

e

$$ax_1^2 + bx_2^2 + cx_3^2 + 2dx_1x_2 + 2ex_1x_3 + 2fx_2x_3 + gx_1 + hx_2 + ix_3 + \alpha = 0.$$

À superficie resultante da equação anterior chama-se **quádrica**. Existem quatro tipos de quádricas não degeneradas): elipsóides, hiperbolóides (de uma ou duas folhas), cones e parabolóides (elípticos ou hiperbólicos).

Exemplo 65. Considere-se a forma quadrática $Q : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$Q(x, y) = 3x^2 + 4xy + 3y^2.$$

Tem-se

$$Q(x, y) = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix},$$

com

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}.$$

Os valores próprios de A são $\lambda_1 = 1$ e $\lambda_2 = 5$. Tem-se então a seguinte forma quadrática diagonal (isto é, sem termos cruzados)

$$Q(x', y') = \begin{bmatrix} x' & y' \end{bmatrix} D \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

com

$$D = PAP^T, \qquad \begin{bmatrix} x' \\ y' \end{bmatrix} = P \begin{bmatrix} x \\ y \end{bmatrix},$$

e

$$P^T = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{4} & \operatorname{sen}\frac{\pi}{4} \\ -\operatorname{sen}\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix}, \quad P = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{4} & -\operatorname{sen}\frac{\pi}{4} \\ \operatorname{sen}\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix},$$

em que P^T é a matriz diagonalizante obtida colocando na 1^a coluna um vector próprio de norma 1 associado ao valor próprio λ_1 e na 2^a coluna um vector próprio de norma 1 associado ao valor próprio λ_2, de tal modo que ambos os vectores próprios constituam uma base ortonormada de $\mathbb{R}^2$. Observe-se que a matriz P é ortogonal, isto é, tem-se $P^T = P^{-1}$.

Tem-se então

$$Q(x,y) = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} =$$

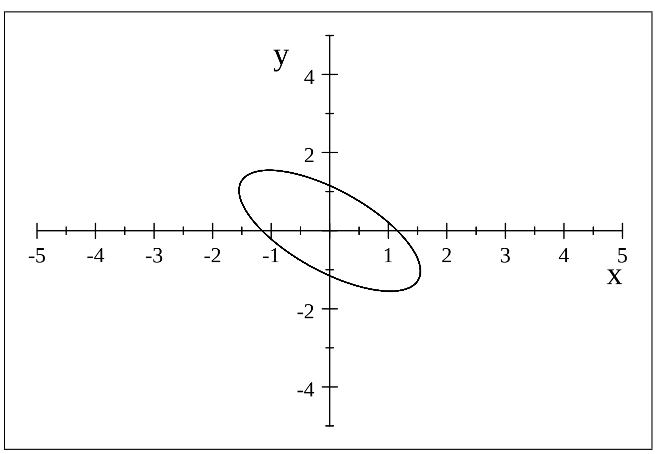
$$\begin{bmatrix} x & y \end{bmatrix} P^T D P \begin{bmatrix} x \\ y \end{bmatrix} =$$

$$= \left(P \begin{bmatrix} x \\ y \end{bmatrix} \right)^T D P \begin{bmatrix} x \\ y \end{bmatrix} =$$

$$= \begin{bmatrix} x' & y' \end{bmatrix} D \begin{bmatrix} x' \\ y' \end{bmatrix} = Q(x', y').$$

Por exemplo, relativamente à equação quadrática

$$3x^2 + 4xy + 3y^2 = 4$$



tem-se a elipse:

$$(x')^2 + 5(y')^2 = 4 \Leftrightarrow \left(\frac{x'}{2}\right)^2 + \left(\frac{y'}{\frac{2\sqrt{5}}{5}}\right)^2 = 1.$$

Definição 87. Seja A uma matriz real simétrica do tipo $n \times n$. Diz-se que A e a forma quadrática $Q : \mathbb{R}^n \to \mathbb{R}$ dada por

$$Q(x) = x^T A x$$

são:

(i) definidas positivas se

$$x^T A x > 0,$$

para todo o $x \neq \mathbf{0}$;

(ii) definidas negativas se

$$x^T A x < 0,$$

para todo o $x \neq \mathbf{0}$;

(iii) semidefinidas positivas se

$$x^T A x \geq 0,$$

para todo o x;

(iv) semidefinidas negativas se

$$x^T A x \leq 0,$$

para todo o x;

(v) indefinidas se existirem pontos onde $x^T A x$ seja positiva e pontos onde $x^T A x$ seja negativa.

Teorema 96. Seja $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ tal que A é simétrica. Então,

(i) A é **definida positiva** se e só se todos os valores próprios de A forem positivos;

(ii) A é **definida negativa** se e só se todos os valores próprios de A forem negativos;

(iii) A é **semidefinida positiva** se e só se todos os valores próprios de A forem não negativos;

(iv) A é **semidefinida negativa** se e só se todos os valores próprios de A forem não positivos;

(v) A é **indefinida** se e só se A tiver pelo menos um valor próprio positivo e outro negativo.

Bibliografia

1. Howard Anton and Robert C. Busby, Contemporary Linear Algebra, John Wiley & Sons, Inc., 2002.
2. Luís Barreira e Clàudia Valls, exercícios de álgebra linear, IST Press, 2011.
3. Maria Esmeralda Sousa Dias, Álgebra Linear,
   https://www.math.tecnico.ulisboa.pt/~edias/TextosNet/ALbookfin_Net.pdf
4. Bernard Kolman, Introductory Linear Algebra with Applications, Prentice Hall, 1996.
5. Steven J. Leon, Linear Algebra with Applications, 8th edition, Pearson, 2009.
6. Seymour Lipschutz, Linear Algebra, Schaum´s Outline Series, 4th edition, McGraw-Hill, 2009.
7. Luis T. Magalhães, Álgebra Linear como Introdução à Matemática Aplicada, 9^a edição, Texto Editora, 2001.
8. Carl D. Meyer, Matrix Analysis and Applied Linear Algebra, SIAM, 2000.
9. António Monteiro e Gonçalo Pinto, Álgebra Linear e Geometria Analítica, McGraw-Hill, 1997.
10. Ana Paula Santana e João Filipe Queiró, Introdução à Álgebra Linear, Gradiva, 2010.
11. Gilbert Strang, Linear Algebra and its Applications, 3rd edition, Thomson Learning, 1988.